APPENDIX E

Peer Review Comments
And
Responses

Summary of Peer Review Comments

The Columbia River Temperature Assessment (Simulation Methods) followed the EPA peer review process as described in EPA (199?). Mr. David Wegner Ecosystem Management International, Inc. and Professor Scott Wells of Portland State University did the formal review. In addition, the review process was opened to other interested parties. The following provided reviews:

Harza Engineering Company under contract to Bonneville Power Administration

Dr. Peter Shanahan of HydroAnalysis, Inc under contract to Potlatch Corporation

Professor M. Bruce Beck under contract to Potlatch Corporation

US Army Corps of Engineering, Water Management Division, Portland Division

Stuart McKenzie

Duane Karna, EPA Region 10, provided colleague review.

Electronic copies were available for all reviews except those done by Stuart McKenzie and Duane Karna. Reviews for which electronic copies were available are attached to the report. Copies of all reviews are on file at EPA Region 10, Seattle Washington (contact Ofelia Erickson at 206.553.1526).

The formal review process included an outline that reviewers could use to focus their review. The results of the review are discussed below in terms of the questions posed in the outline.

Part I. Conceptual Model

(1) Have the objectives of the temperature model been clearly identified?

A number of reviewers had trouble identifying the objectives of the assessment. Many concluded (incorrectly) that the objective was to develop policy regarding specific management strategies for attaining water quality standards and restoring endangered stocks of salmon and steelhead. In the final report, the section on objectives was expanded and revised in an effort to clarify the objectives of the temperature assessment, which were simply to assess the relative impacts of dams and tributaries on the temperature regime of the Columbia and Snake rivers

(2) Has the level of certainty required by the model objectives been identified and can the proposed concept achieve this level of certainty?

The reviewers who responded directly to this question concluded the level of certainty identified in the analysis addressed the objectives for developing a screening model. The term "screening model" has been removed from the report. The objective of the analysis was to assess the relative impact of dams and tributaries on the temperature regime of the Columbia and Snake rivers. This objective implies that further study of both the modeling system and the monitoring program may be necessary prior to using the model for policy-making. An important part of such a study would be the development of a protocol for determining the level of certainty required by the modeling and monitoring system.

(3) Have the appropriate system boundaries, time scales, and length scales been identified?

In response to comments from some reviewers that the geographic scope was not adequate, the model was extended upstream on the Snake River to R.M. 168. And the Clearwater River was added as a tributary. This increase in geographical scope does not address all the sources of temperature modification in the Columbia and Snake rivers. However, it does make it possible to evaluate control strategies related to the operation of Dworshak Dam on the North Fork of the Clearwater River.

(4) Have the important source terms and background conditions been identified and are there adequate data to characterize them sufficiently for the model application?

Reviewers raised a number of questions regarding the approach used to develop input temperatures for major tributaries and for extending meteorological data over distances of 60-100 km. The problem of extending data in both time and space is a universal one, associated with the lack of resources needed to obtain all the necessary data. The Report recognizes the uncertainty associated with the space and time gaps in data and uses peer-reviewed methods or widely used assumptions to address this uncertainty.

(5) Are the available data adequate for achieving the levels of certainty required by the model objectives?

Reviewers agreed the data were adequate for purposes of a temperature assessment, but mentioned the need for better meteorological data and input temperatures from tributaries and groundwater. In addition, one reviewer was highly critical of the temperature monitoring program at hydroelectric facilities on the Columbia and Snake rivers. The data review of McKenzie and Laenen (1998) also implies that better quality assurance and monitoring design are needed for the temperature monitoring program.

Part II. Model Development

(1) Is the model being developed based on current knowledge and do the mathematical descriptions accurately reflect the processes identified in the conceptual model?

Most reviewers agreed the model was based on current knowledge. One reviewer, however, was critical of the approach used to develop the thermal energy budget at the air-water interface. This reviewer proposed an alternative method they had developed for other river systems. However, this method did not appear to be peer-reviewed, nor was it made available for use in the temperature assessment.

(2) What structural properties in the model could affect reliability of model predictions? (3) Is the parameter estimation process reasonable in terms of available data and knowledge?

Reviewers discussed several aspects of structural properties of the model. This included a caution that the model should not be used for periods of extremely unsteady flow (snowmelt and storm conditions) because the hydraulic model is a steady-state gradually-varied flow model. As mentioned previously, one reviewer felt the heat budget method contained flaws and proposed a method that has been neither peer-reviewed nor made available for use by others. One reviewer questioned the structure and use of the uncertainty propagation scheme, although this reviewer agreed the uncertainty propagation scheme had followed peer-reviewed methods. Finally, a reviewer raised the question of whether or not the computer code itself had been adequately tested

The uncertainty associated with the gradually-varied flow hydraulic model is discussed in the Report. Peer-reviewed methods were used to develop the heat budget and the uncertainty propagation scheme. An extensive discussion of tests of the computer code was added to the Report in response to the comment regarding adequacy of the code.

(3) Is there are well-designed plan for determining if and when the model is acceptable for use as a decision support tool?

Reviewers did not comment on this directly, though in general they believed there should be more discussion of the differences in observed and simulated results. In response, the discussion of uncertainty was expanded and a number of statistical analyses were added to the Report.

(5) Are all the components of the conceptual model realized in the model development?

In general, reviewers concluded that the components of the conceptual model were realized. One reviewer felt there should be more discussion of reservoir limnology and hydrologic dynamics. Some discussion of temperature impacts on salmon and steelhead was added to the Report, but an expanded discussion of these issues was deferred to the next phase of the TMDL process. The next phase of the TMDL process is the development of a problem assessment and establishing priorities for management strategies.

Part III

(1) Do the model results adequately address all the objectives?

Reviewers concluded the model addressed the objectives for a temperature assessment.

(2) Do the results properly characterize the uncertainty and variability associated with data collection, source characteristics and model error?

A number of reviewers raised questions regarding uncertainty and variability associated with data collection source characteristics and model error. The discussion of uncertainty and variability was expanded considerably in the Report. However, many of these issues were left unresolved simply because there is, at present, no specific protocol for evaluating model uncertainty and variability. The work described in the Report represents a much more comprehensive approach to the issue of uncertainty and variability than previous studies of the Columbia and Snake rivers. Nevertheless, this general topic is one deserving of more attention by both scientists and policy-makers.

(3) Are the conclusions reasonable in terms of the model/data uncertainty and variability?

The formal reviewers stated the conclusions were reasonable for the stated objectives of the temperature assessment.

(4) Is the work documented well enough such that others could reproduce the results?

The only reviewer who responded to this question concluded the work could be reproduced assuming responses to comments were adequate. An effort was made to respond to all reviewers' comments. However, comments related to broader watershed issues and more thorough limnological studies were considered to be beyond the scope of the analysis. That is not to deny the importance of these topics, rather to defer them to that part of the water quality planning process which includes problem assessment and development of strategies for attaining water quality standards.

Editorial/Style Comments

Most comments of an editorial nature were either incorporated directly into the report or used to modify or expand the text. Some comments regarding style were incorporated into the Report. However, when the comments did not appear to be consistent with the scope of this phase of the study, they were not included in the Report.

Memorandum

May 24, 1999

To: John Yearsley, EPA

From: Scott Wells, PSU, scott@eas.pdx.edu, (503)725-4276

Re: Peer Review comments on Temperature Modeling in Columbia River basin

Based on our phone conversation of 5/17/99 and your letter of 4/23/99 in response to my draft comments, I have enclosed a revised set of comments on the report entitled: "Columbia River Temperature Assessment: Simulation Methods."

Please contact me for further clarification on these comments. I have tried to be objective in evaluating the report you provided. Some of my comments may be based on only having partial understanding of the issues and modeling approach. Any further clarification would be useful for me. Thank you for the opportunity to work with you on evaluating this modeling effort.

Review of "Columbia River Temperature Assessment: Simulation Methods" by J. Yearsely, EPA

Modeling the Columbia River is a difficult and complex undertaking. Dr. Yearsley is to be commended for approaching this task. Modeling such large systems has many challenges from modeling to data analysis. The review comments are divided into minor typographical and more substantive comments and critique. Two additional questions posed by Dr. Yearsely in a letter dated 4/23/99 were also evaluated.

Minor Typographical Comments

1. p. 3 under "Hydrology" 3rd sentence change "can longer" to "can no longer"

Response: Text was modified as suggested

2. p. 3 under "Water Resources Development" 2nd paragraph 2nd sentence change "The systems has" to "The systems have"

Response: Text was modified as suggested

3. p. 5 2nd paragraph eliminate reference to "(Barnwell and Krenkel, 1982)" since it is redundant.

Response: Text was modified as suggested

4. p. 7 1st paragraph change "(Cole and Buchak (1995)" to "(Cole and Buchak, 1995)"

Response: Text was modified as suggested

5. p. 9 next to last paragraph, last sentence. Change "...such as dispersion and turbulent diffusion" to "...such as dispersion" since in a 1-D system turbulent dispersion is always >> turbulent diffusion.

Response: Text was modified as suggested

6. p. 12 last sentence on page requires a period

Response: Text was modified as suggested

7. Table 5: Data Source for 1st row reference is missing date

Response: Text was modified as suggested

8. p. 20 1st paragraph last sentence change "Bonnevill" to "Bonneville"

Response: Text was modified as suggested

9. p. 20 last paragraph change "xceedance" to "exceedance"

Response: The word "exceedance" was removed from the report

10. Figures 22-28: cannot tell from legend which graph is which

Response: Text was modified as suggested

11. Figure 26: y-axis title is incorrect; it should be variance not temperature

Response: Text was modified as suggested

<u>**Comments/Critique**</u>

1. p. 4 Water Quality Issues

The 303(d) list for the State of Washington is used. Does this agree with the State of Oregon 303 (d) list for the Columbia River ? Perhaps a comment is required here to address State of Oregon 303 (d) issues, if there are any.

Response: Text was modified to include discussion of the Oregon DEQ 303 (d) list of water quality segments.

2. p. 7 Thermal Energy Budget

The assumptions in the development of Eq. 1 should be clearly stated as (1) no dispersion and (2) cross-sectional homogeneity (implying that all inflows mix laterally in the reach they are added).

- Is the assumption of no dispersion justified? Since the model uses steady-state hydraulics and computes only daily averaged temperature (as mentioned later in the report), dispersion may not be important. A calculation could show that dispersion is not an issue. This can be done by computing the Peclet number, including the effect of heat transfer.

Response: A quantitative discussion of dispersion has been included.

- Since the temperature standards are usually written in terms of daily maximum temperatures, the inclusion of dispersion in the temperature model will make a difference in the prediction of daily maximum temperatures or instantaneous temperatures. Dispersion therefore should be part of the model if the model were being used to evaluate instantaneous temperature standards in the Columbia River.

Response: The benchmark of 20 °C was used to assess water temperature impacts. A discussion of the rationale for choosing this benchmark has been added to the report

3. p. 8 Eq. 3 Define the term T_k

Response: Text was modified as suggested

4. p. 8 Eq. 4. Make sure that H_{evap} (also p.15 Eq. 19) is constrained always to be "evaporation" and not "condensation". Using the formula of Eq. 19 one can obtain conditions where H_{evap} adds energy to the water body. This should never occur, and the model needs to be constarined that H_{evap} is only a loss of energy or zero.

Response: The model includes this constraint.

5. p. 10 The choice of the mixed Eulerian-Lagrangian method is appropriate for this system. This reviewer does not see advantages to using this technique over other techniques since interpolating temperatures using a 2nd order polynomial introduces "diffusive" error into the Lagranian technique that was trying to be avoided. Hence, there is no clear superiority of the Lagrangian technique over a purely Eulerian technique that has the same order of spatial accuracy.

Response: A discussion of the advantages of the mixed Eulerian-Lagrangian method is included in the report. In addition, an analysis of the numerical error associated with the reverse particle tracking scheme is compared to that of two numerical methods, WQRRS and QUAL2E, that have been used to assess water temperature in the Columbia and Snake Rivers. For the test conditions analyzed, reverse particle tracking was clearly superior toWQRRS and QUAL2E in terms of its capability to propagate high frequencies.

6. p. 10 The "geometric properties of the river system" were assumed to be constant during a given time step. It was unclear whether this was the numerical time step or the "daily-averaged" temperature model period. If this assumption was for a daily time step, there could be large errors in the water balance if the time step were this coarse.

Response: The geometric properties of the river system were assumed to be constant for the duration of one day. Since the dams included in this analysis are operated as run-of-the-river , the assumption would seem to be a reasonable one. For the case with dams removed, some error may be introduced. However, since the flow at USGS gages is reported on a daily basis, it is difficult to make a quantitative estimate of this error.

7. p. 12 Time and Length Scales

A discussion was made about the time and length scales of the forcing functions of the system. These can be determined explicitly from data sets (especially the 3-hourly meteorological data) using Spectral Analysis where the data are evaluated using Fourier transforms. Not only are the important frequencies determined from this type of analysis, but their relative importance is also determined.

Response: The highest frequency included in the assessment was associated with the daily average. This frequency was consistent with the objectives of the assessment and with the frequency response of the numerical scheme.

8. p. 15 Heat Budget

The heat budget terms are properly formulated, but Eqs. 18-21 require the definition of several more terms before these can be evaluated, such as the terms λ, E_v, T_{DB}, W, e_s, e_o, e_a, etc.

Response: Text was modified as suggested

9. p. 15 and 16 Initial Water Temperatures

A regression equation, Eq. 22, was used to compute stream temperatures based on air temperatures. With the 4 parameter model of Eq. 22, statistical curve fits probably are very good. A more accurate correlation though may be based on equilibrium temperatures rather than air temperatures. Using air or equilibrium temperatures may not affect the accuracy of the model. The comment that the stream temperatures predicted by Eq. 22 gave "good results even when the air temperature measurements were not in proximity to the stream gaging station locations" is an indication that the correlation is not a strong function of local air temperature but is only a result of calibration to a 4 parameter statistical model.

Also, the correlation is based on "weekly stream temperature" even though the in-stream model is "daily averaged" and the TMDL requirements for temperature are based on instantaneous daily maximums. The use of a weekly stream temperature was not explained in the report. Why were daily average temperatures not used in order to be compatible with the in-stream model temporal resolution?

Response: The method used to estimate stream temperatures was based on a peer reviewed study.

10. p. 17 Systems Model Bias and Error

The statement that "wind speed, cloud cover, relative humidity and station pressure are large-scale phenomena and that air temperature is a more local phenomenon" is not necessarily correct. I understand in the modeling of such a large system the modeler has to make approximations, but the above statement is not a good basis for ignoring local variability in wind, clouds, and humidity. An analysis showing the meteorological variability and the expected degree of error induced by neglecting that variability would be useful.

Response: The assumptions regarding meteorological data were not meant to be the basis for ignoring local variability. Rather they were meant to describe the process used in many temperature studies (references provided in the text of the report) to account for the limited availability of the necessary meteorological variables. A discussion of regional correlation between meteorological variables has also been added to the text .

Using only 3 meteorological stations with high-resolution data does not seem to provide the required spatial reliability for such a large system. Echoing the comment above, a special effort should be undertaken to provide a much more rigorous analysis of meteorological data variability in the basin. Such information was not available for review in the report.

Response: Model uncertainty, including that associated with meteorology, is lumped into a single noise term.

11. p. 19 Uncertainty and Variability

The purpose for this modeling exercise was to "identify critical issues for additional study". The peer review process is serving that purpose.

The statement was made that "the focus in this study was on the space-time complexity rather than on model complexity". All modeling should be done at the appropriate level of complexity for the project – otherwise the usefulness of the model study is compromised. If the model used is not appropriate or does not provide enough detail to use for TMDL analyses, a more complex model should be used. This modeling objective and purpose may need to be included earlier in the paper.

Response: As described in the report, the level of model complexity was appropriate for the stated objectives and equal to, or greater, than similar studies used as decision support tools for policy-making in the Columbia and Snake rivers.. The level of complexity associated with spatial and temporal scales was higher than for previous studies of water temperature to achieve the objectives of characterizing the uncertainty and variability of the system.

Comments on each model component:

- Heat budget: components of the heat balance are reasonable but the model is adequate only for steady-state simulations since dispersion is neglected. There is some numerical "diffusion" because the Lagrangian solution is interpolated onto a grid. The model is not adequate for instantaneous temperature simulation.

Response: Instantaneous temperatures were not simulated. A daily averaged water temperature was selected as the benchmark..

- River hydraulics: model is adequate for steady-state hydraulics. Hence, for periods of unsteady-flows (snowmelt and storm conditions), the model is inadequate.

Response: Comment noted

- Initial conditions – reasonable especially since they do not affect conditions long into the simulation

Response: Comment noted

- Water Balance – Even though the conclusion is made that irrigation, groundwater return flow and miscellaneous tributary flows were only 5-7% of the flow increment in the Columbia, there is no reason not to include these if they are know since they probably have an important local effect.

Response: Effects of groundwater return and the Walla Walla River were added to the analysis

- Filter – The discussion in the text was unclear as to how the Kalman Filter technique was used. Initially, this reviewer thought that the Kalman Filter was used in predicting the new state of the Columbia

River without dams while using statistical estimates based on the existing system. In discussions with Dr. Yearsely, this was clarified. Apparently, the Filter was used only to estimate the variance of the estimate of temperatures from the deterministic model. In this sense, the Filter is appropriate to use. Apparently, figures 6-13 are the results of the deterministic model and not the results after using the Kalman filter. This should be clarified in the text.

Response: Additional discussion was added to the report

Also, since many are familiar with Monte Carlo techniques, the text should at least mention why Kalman Filter techniques were chosen over Monte Carlo techniques for assessing the variance of the deterministic model estimate.

Response: The report was modified as suggested

- The deterministic model as a whole (Heat balance equation including heat budget and river hydraulics): The deterministic model does not appear to be accurate during transition seasons. There needs to be statistics showing the mean error of the model at each location in figures 6-13. In general, a deterministic unsteady temperature model without any statistical filtering should have a mean error less than 1 and 2°C for instantaneous temperature predictions. If instantaneous temperatures are required to be predicted, a more complex model may be necessary.

Response: Instantaneous temperatures were not modeled.

12. In Figures 6-13: Temperature data in the winter approach 0°C even at Bonneville Dam in the winters of 1991 and 1993. Is this realistic?

13. In the graphs showing "Frequency of Exceedance" Figures 30-41, comparisons of graphs with and without dams and with tributary temperatures altered need to be on the same graph rather than on separate pages. This would aid the reader in evaluating comparisons between alternatives.

Response: Comment noted

14. In the results section, much of the discussion relates to the average magnitude of exceedances of the management strategies. This is appropriate. But it should also be stated in the text that the temperature standard deviations of the alternatives many times overlap such that it is possible that there may be no difference between management strategies.

ext was modified as suggested

n of Model Objectives and Model Certainty

In a ... er of 4/23/99, Dr. Yearsely asked that 2 questions be evaluated. The questions and the evaluations are included below.

1. Have the objectives of the temperature model been clearly identified?

The model objectives were mentioned briefly on the bottom of p.1, end of p.4 and top of p. 5, and again on p. 18. It should probably also be stated in the report that this work is not meant to look specifically at the temperature TMDL issue. This becomes a little unclear when there is a discussion on the temperature TMDL on p.4 followed by the statement at the bottom of p. 4 that this is "a first step in developing a TMDL." If one were developing a model for the TMDL evaluation, the model chosen would be somewhat different as explained above in specific review comments.

2. Has the level of certainty required by the model objectives been identified and can the proposed concept achieve this level of certainty?

The level of certainty required to achieve the model objectives was not explicitly stated in the report and is a very subjective measure. This would be more easily evaluated if the model deterministic error were clearly shown relative to the expected differences predicted by the model for different alternatives. In general, though, the level of certainty of this work is consistent with the objectives of a screening model analysis.

ECOSYSYSTEM MANAGEMENT INTERNATIONAL, INC.
2609 COLUMBINE AVENUE
DURANGO, CO 81301
(970) 259-2510
EMAIL: emiwegner@aol.com

March 31, 1999

To: Mr. John Yearsley, Environmental Protection Agency
Region 10
1200 Sixth Avenue
Seattle, Washington

From: Dave Wegner
Ecosystem Management International, Inc.
Principal Scientist

Subject: Review Comments of Technical Report
Columbia River Temperature Assessment: Simulation Methods

I have received and reviewed the technical report entitled ***Columbia River Temperature Assessment: Simulation Methods***. I have reviewed the documentation that you provided and have evaluated your approach, assumptions and results based on my experience and historic application of thermal models in the Colorado River basin.

My review follows the following outline and represents the major areas that your report addresses. My review approach has focused on three primary representations.

- Conceptual Representation - has the logic for development of the model been adequately laid out and are the steps for application clearly defined?
- Functional Representation - has the formulation of the model, specifically the physical constraints, process, variables, and boundary conditions been adequately defined?
- Computational Representation - does the model adequately translate the logic into correct mathematical forms and procedures necessary for solution of the problem over the desired temporal and spatial spectrum?

My primary expertise lies in the evaluation of the Conceptual and Functional representation arenas. These two areas must be credibly and accurately defined if an accurate assessment is to be completed on the thermal conditions in the Snake and Columbia Rivers.

MODEL PEER REVIEW FOCUS QUESTIONS

Part I. Conceptual Model

(1) Have the objectives of the temperature model been clearly identified?

In general yes. Additional clarification is needed in regards to who will be using this model, what level of detail is required in their use, and if this model will be used to set TMDL limits.

(2) Has the level of certainty required by the model objectives been identified and can the proposed concept achieve this level of certainty?

In regards to the objective of developing a screening model, this assessment has achieved its goal. Applications above the screening level however require additional clarification, statistical analysis and a more rigorous assessment of the error bias.

(3) Have the appropriate system boundaries, time scales and length scales been identified?

For a screening model assessment the appropriate boundaries and scales have been identified in general except for inclusion of addressing the boundaries related to reservoir dynamics behind the study dams. This area needs to be expanded upon (see comments below).

(4) Have the important source terms and background conditions been identified and are there adequate data to characterize them sufficiently for the model application?

In general yes. The assessment does a good job of identifying the necessary model parameters and the data necessary for application. The boundary conditions for the reservoir and pre-project thermal and flow conditions however should be further articulated.

(5) Are the available data adequate for achieving the levels of certainty required by the model objectives?

For a screening level model the answer is yes.

Part II. Model Development

(1) Is the model being developed based on current knowledge and do the mathematical descriptions accurately reflect the processes identified in the conceptual model?

Yes, the model is based on the present state of the art.

(2) What structural properties in the model cold affect reliability of model predictions?

- Reservoir limnology is not dealt with directly and should be.
- The physical geometry needs to be beefed up for any level of analysis beyond the screening level.
- System, both physical and limnological, variability is not addressed adequately for anything beyond a screening level approach.
- The statistical significance of the results has not been addressed and should be before the report is finalized.
- Temporal and spatial variability should be discussed as related to implications to model results

(3) Is the parameter estimation process reasonable in terms of available data and knowledge?

For a screening level assessment the answer is Yes.

(4) Is there a well designed plan for determining if and when the model is acceptable for use as a decision-support tool?

A rigorous statistical evaluation of the results is necessary before providing this tool to decision-makers. Additionally an assessment on how this model should be used must be developed before it is put on the street. There is not a plan presented in the report and there should be.

(5) Are all components of the conceptual model realized in the model development?

As developed for a screening process the answer is Yes. However there are several critical areas that should be discussed and evaluated prior to further application:

- Inclusion of reservoir and hydrologic dynamics
- Inclusion of a statistical analysis of results
- Clarification of model application. The parameters, especially the innovation process needs to be explained further
- The model approach has been well done for the level of answer desired. My fear is that the public will rush to conclusions without fully understanding the constraints necessary in interpreting the results or that this was a screening level assessment.

Part III.

(1) Do the model results adequately address all the objectives?

The results address the objective of developing a screening model. The results do not address the cause for the increase temperatures (dams, watershed development, or hydrologic modification). It is clear that something has raised the temperature and it is probably in a priority based on: (1) dams; (2) changes in regimes; (3) watershed impacts. Care should be taken with a screening level study of jumping too far out on the limb.

(2) Do the results properly characterize the uncertainty and variability associated with data collection, source characteristics a model error?

At a screening level model the answer is yes. Additional work should be done however to address the impacts related to the reservoirs, retention time, and statistical significance of the differences in the results.

(3) Are the conclusions reasonable in terms of the model and data uncertainty and variability?

Yes with the caveat that a statistical evaluation of the significance of the difference as related to the model error should be made before the conclusions are made public.

(4) Is the work documented well enough such that others could reproduce the results?

Yes IF the comments made below are addressed.

Overall I feel that the approach, methodology and application was very well done for the level of analysis described. My specific comments included below represent my review of the document and my attempt to make the document more readable to the interested public and decision-makers. My hope is that the document will reach the necessary managers and result in support for expansion of the study to include a more rigorous evaluation of the model, the variability of the system and the application to additional alternatives.

SPECIFIC COMMENTS ON THE REPORT

I. Introduction and Background

The objectives of the report is stated *to assess the relative importance of different sources of watershed impact in respect to changes in the temperature regime of the main stem Columbia River in Washington and Oregon and in the Snake River in Washington.* Three general sources of river impact are identified:

1. Construction of impoundments for hydroelectric facilities and navigational locks.
2. Hydrologic modification to the natural river system as related to irrigation and navigational development
3. Modification of the watershed from agricultural and silviculture practices which reduce riparian vegetation, increase sediment loads and change stream or river geometry.

Ultimately the model will be used by managers and decision-makers to evaluate a decision support system for developing management strategies for attain water quality standards and protect beneficial water uses.

<u>Comments:</u>

- A more complete identification of why this modeling approach is being used as related to the three impacts is necessary.

Response: A discussion of the rationale for developing and applying the model to the Columbia and Snake rivers is described in the Report

- How was the decision made to develop this model? Was this an EPA directive? A request from National Marine Fisheries Service (NMFS)?

Response: The model was develop by EPA Region 10 as part of the planning process in Section 303 (d) of the Clean Water Act.

- The objective of this assessment should be made in the introductory section

Response: The goals of the assessment are stated in the Introduction and a section entitled, Study Objectives, has been added to the Report.

- Consider adding a short glossary of important terms

Response: Suggestion noted

II. GEOGRAPHY, CLIMATE AND HYDROLOGY OF THE COLUMBIA BASIN

Comments:

- One of the sources of impact that this assessment is to address is watershed development. In that case, a more definitive evaluation of the watershed that may affect the water temperature of the study area should be identified.

Response: This analysis focussed on the main stem Columbia and Snake rivers. An analysis of the watershed was beyond the scope of this assessment.

- Reference is made to *confounding tributaries*. Where are these tributaries and what are the seasonal influence on the overall river thermal integrity?

Response: The tributaries included in the analysis are described in the report as are the sources of temperature data for characterizing seasonal variations.

- Can you prioritize which tributaries in the supporting watershed have the most potential for impacting the results of the assessment? Percentage or location wise which ones need to be concerned about?

Response: The analysis of the watershed was beyond the scope of this report. Some additional discussion has been provided in the report elaborating on the relative importance of various tributaries.

- The pre-project hydrology should be identified. This should be in two leve
  a. Seasonal (monthly) perspective
  b. Daily regime (how much daily fluctuation occurred?

Response: Pre-project conditions were not part of the analysis. The objective of the assessment was to evaluate the relative impact of dams and tributaries on water temperature given existing management of the system and variability in meteorology and hydrology represented by the period 1975-1995.

- The post project hydrology regime for high, average and low water years should be presented. This would provide a spectrum of what the hydrologic boundaries. This is important in regards to evaluating the model.

Response: Hydrologic data from USGS gaging stations on the Columbia and Snake rivers for the period 1975-1995 were used for the analysis. Data sources have been referenced in the Report.

- The seasonal and summer/monthly flow regimes should be identified for the management of the dam complex. This is important in regards to interpreting when the thermal thresholds occurred and how well the model predicted reaching the thresholds.

Response: See comment above regarding source of hydrologic data.

- An idealized hydrologic regime should be presented for each of the alternatives that the model is expected to be used to evaluate. In this way it can be determined how well the model is matching predicted flow scenarios.

Response: Flows were not predicted in this analysis. See comment above regarding source of hydrologic data.

III. WATER RESOURCES DEVELOPMENT

This section of the report needs to be significantly expanded upon to address not only the time of development of the four dams but also what this has meant to the hydrologic and therefore thermal regime of the study area. The broad sense of the Columbia Basin development is addressed adequately however the specific relationship to the project area needs to be discussed.

Comments:

- Expand upon the development of the four study area dams. What impact did they have on the pre-project flow regime?

Response: All the dams on the Columbia River below Grand Coulee Dam and all the dams on the Snake River below Lewiston, Idaho were included in the Report. Pre-project conditions were not evaluated in this study.

- Discuss
- how the dams are operated. Are they operated as run-of-the-river, periodic storage, flood control, navigation, stabilization for downstream releases? Where is water withdrawn at the dams?

Response: A discussion of dam operation, commensurate with the scope of the analysis, has been provided.

- Discuss the physical and limnological effect of water resource development in the project area. Specifically add a section on the limnological relationships that occur as a result of flow regulation. Percentage of the time that stratification occurs? What is the residence time of water within the reservoirs?

Response: The effects of water resource development will be discussed in a problem assessment of the Columbia and Snake rivers.

- What are the upstream impacts as related to Hells Canyon dam releases? Does it have a seasonal warming effect? What are the input conditions and does it affect the thermal capacity of the study area?

Response: The geographical scope of the analysis is limited to the Columbia River below Grand Coulee Dam and the Snake River below its confluence with the Grande Ronde River. Some of the effects of upstream management can be inferred from the magnitude and frequency with which temperatures exceed the benchmark at the upstream boundary.

- Are there any impoundments on the tributaries that may be confounding the problem by providing seasonally warmer water?

Response: This analysis focussed on the main stem Columbia and Snake rivers. An analysis of the watershed was beyond the scope of this assessment.

- A *Biological Relationship* section should be added that identifies the critical biological threshold levels of the primary species of concern in the study area. This is important in that it provides a frame of reference in regards to evaluating the assessment. If the predictions are close to the threshold it is

utting forth-additional efforts to fine-tune the numbers. If the
ons are not within the proximity of the threshold then not as much
ay be required to assess the trend or direction of the prediction.

A discussion of the biological effects of temperature on salmonids has been added to the Report.

- What was the pre-project thermal profile for the river within the study area?

Response: Pre-project conditions were not part of the analysis. The objective of the assessment was to evaluate the relative impact of dams and tributaries on water temperature given existing management of the system and variability in meteorology and hydrology represented by the period 1975-1995.

- How was the 20-degree Centigrade level identified? Reference where this came from and what it is supposed to protect. Salmonids? Macroinvertebrates? Humans?

Response: A discussion of the rational for choosing 20 °C as the benchmark and a discussion of the effects of temperature on salmonids has been added to the Report.

IV. STUDY OBJECTIVES

Comments:

- On Page 5 it is stated that the purpose of TMDL assessment is to:
  - Identify the sources of water quality parameters of concern
  - Identify what if any control or management strategies are possible

It is stated that the temperature assessment models will be used to provide some of the framework for a problem assessment in the mainstem Columbia River. *Is this model ultimately going to be used for the development of the TMDL? If so, how was it determined that this was the best model for use?*

Response: EPA and the states of Oregon and Washington have not yet determined how a TMDL will be performed on the Columbia and Snake rivers. A discussion of the rationale for choosing the model has been provided in the Report.

- The <u>objective</u> of the assessment is defined as being to *develop and implement a mathematical model of water temperature for the Columbia and Snake Rivers in a way that is generally consistent with those of the screening model.* That stated, what is the level of detail that is required to address the questions being asked? In other words it should be stated how good the model has to be - within one degree? One level of statistical significance? Etc. The point is that it should be stated what the expectations and requirements are so that we can adequately determine if the model is meeting those objectives.

Response: The level of significance necessary for subsequent planning and decision-making would be determined by the programs responsible for watershed planning. The uncertainty analysis provided in the assessment provides a basis of determining the level of significance or risk associated with using the model as a decision support tool.

- Are there any biological or engineering objectives in this assessment?

Response: The objective of the assessment was to evaluate the relative impact of dams and tributaries on water temperature given existing management of the system and variability in meteorology and hydrology represented by the period 1975-1995. Biological and engineering objectives would be part of the watershed planning process.

V. MATHEMATICAL MODEL DEVELOPMENT

This section has five (5) sections. Comments will be separated into the appropriate section.

<u>System Boundaries</u>

- No mention is made of the four reservoirs within the study area and the boundaries associated with them.

Response: Characteristics of the reservoirs on the Columbia and Snake rivers are provided in the Report.

- Are the tributaries included within the watershed system boundary?

Response: A description of the tributaries included in the analysis are provided in the Report.

- What are the hydrologic system boundaries associated with this assessment?

Response: Hydrologic boundaries correspond to the thermal energy input boundaries. Tables describing the location of these boundaries are provided in the Report.

- Figures similar to the "Surface elevations in Lake Franklin D. Roosevelt du 1998" should be made for each of the four reservoirs in the study area ove range of hydrologic regimes. This would help to identify the impacts of flow to the transfer of heat energy.

Response: Characteristics of the reservoirs on the Columbia and Snake rivers are provided in the Report.

- Was 1998 a "typical" year hydrologically and thermally at Lake FDR?

Response: Surface water elevations for Lake FDR for 1998 provided an example of typical excursions in the volume of Lake FDR compared to those in the run-of-the-river reservoirs. While these time series may not be "typical" they are "representative" of the way run-of-the-river reservoirs are operated compared to the way Lake FDR is operated

- The present baseline boundaries need to be identified for upstream and downstream positions on a seasonal basis.

Response: A discussion of the geographic scope of the temperature assessment has been provided in the Report.

Thermal Energy Budget

- The statement of *The thermal energy budget has proven to be a useful concept for simulating.....* needs to be referenced. Who has proven it?

Response: A discussion of other applications of the thermal energy budget has been provided.

- Have studies been done using the Eulerian approach rather than the Lagrangian approach? Where? How successful?

Response: A discussion of various approaches to numerical modeling of water quality in rivers and reservoirs has been provided.

- How are reservoir impacts accounted for in this approach?

Response: Reservoir impacts are accounted for in this analysis primarily by the change in system geometry and by constraining the elevations of the reservoirs to be constant throughout the year.

Solution Method

- What is a likely range of the Kalman gain matrix-weighting factor? Do weighting factors connote large potential errors in evaluating the results the assessment?

Response: The Kalman gain matrix can vary from 0 to 1. A value of implies that when making an estimate of the system all the weight is given to the systems model. A value of 1 implies that all the weight is given to the measurement. For values between 0 and 1 weight is distributed between systems model and measurement model according to their relative variance.

- Define the Courant stability criterion (page 10)

Response: A definition of the Courant stability criterion has been added to the Report.

- It is stated on page 10 that the mixed Eulerian-Lagrangian method is used in the models. Once the river was subdivided into "N" segments for analysis was any validation done to check to see if the spatial segments provided the constant thermal properties necessary for the solution approach? In other words, once the model time and spatial steps were determined was there any work completed to determine if those assumptions were indeed correct?

Response: The only test of the assumptions was to compare the simulated water temperatures and the observed water temperatures.

- Can a flow diagram of the sequence of operations performed in the solution of the thermal equations be provided?

Response: Comment noted

<u>Time and Length Scales</u>

- Pre-project (development) hydrologic and thermal regimes need to be included in this analysis in order to ascertain the correct time and length scales.
-

Response: Pre-project conditions were not part of the analysis. The objective of the assessment was to evaluate the relative impact of dams and tributaries on water temperature given existing management of the system and variability in meteorology and hydrology represented by the period 1975-1995.

- Was a statistical analysis completed (with the existing data) to determine the variability of the pre/post project regimes? This would assist in determining the time and length step required.

Response: A statistical analysis that is appropriate for the scope of the analysis has been added.

- Is this model only going to be used to evaluate existing operations? Will there not be a need to determine what could be done if the alternative to breach the dams is evaluated?

Response: The objective of the assessment was to evaluate the relative impact of dams and tributaries on water temperature given existing management of the system and variability in meteorology and hydrology represented by the period 1975-1995. No management options, including breaching of any of the dams, were analyzed in this Report.

- What is the source of the geometric data? What is the stream channel variability?

Response: Sources for the channel geometry are provided in the Report. The coefficients used to characterize channel geometry as a function of flow are also provided.

Rationale for Approach

- Have any of the approaches identified on Page 12 gone through review to the level that the conclusion to use the mixed Lagrangian-Eulerian scheme is adequate for the quality of answer needed in this assessment?

Response: A discussion of various approaches to modeling surface water quality is provided in the report.

- Since it appears that development of the TMDL is a primary goal of this assessment, has EPA defined/recommended the level of detail required?

Response: The objective of the assessment was to evaluate the relative impact of dams and tributaries on water temperature given existing management of the system and variability in meteorology and hydrology represented by the period 1975-1995. Decisions to conduct further watershed planning will be made by the program offices

- Have other models been evaluated as potentially appropriate to this assessment?

Response: A discussion of other surface water quality models and the rationale for model selection is provided in the Report.

- Does the level of effort in this model match the level of quality required for the decision-makers?

Response: EPA Region 10 believes the level of effort devoted to the model matches the quality required by decision-makers.

VI. DATA SOURCES

Comments:

- Is the quality of the tributary data consistent with the quality of the thermal data compiled by Laenen and McKenzie, 1998?

Response: Quality of the tributary data, at least in terms of temporal coverage, is somewhat poorer than thermal data compiles by Laenen and McKenzie. Tributary data are generally collected on a weekly or monthly basis. Therefore, it was necessary to use interpolation methods based on local air temperatures to interpolate weekly or monthly observations to daily tributary temperatures. As noted in the report, this introduces some uncertainty into the final result. The magnitude of the uncertainty is related to the relative contribution of the tributary to the thermal energy budget of the main stems.

- Is the thermal data spatially distributed adequately to allow for model evaluation? In other words are there thermal sampling points at locations where the model will be making intermediate predictions?

Response: The thermal data, though at times of questionable quality, is spatially distributed adequately to allow for model evaluation.

- How as the information in Table 5 consolidated for use in the model? Were representative sections used or were specific hydrologically important locations selected?

Response: Geometric data was selected so as to provide spatial coverarage at a scale of one to ten miles.

- Was channel roughness considered in the development of the model?

Response: The steady-state gradually varied flow model, HEC-RAS, was used to describe system hydraulics as a function of flow. Channel roughness is a required input to HEC-RAS.

- Are the gaging stations adequately spaced?

Response: The gaging stations are adequately spaced

- Is solar radiation important to the heat transfer evaluation? If so, was there any solar information collected?

Response: Solar radiation is simulated using peer reviewed methods described in the Report.

- Was time of water being impounded behind the dams considered in the assessment? What is the retention time of the reservoirs and is there any indication that seasonal, daily of vertical stratification occurs?

Response: Hydraulics of the river in both the impounded and unimpounded condition were analyzed. System hydraulics play an important role in the thermal energy budget.

VII. PARAMETER ESTIMATION

<u>Comments:</u>

Deterministic Elements = Source term = heat budget + advected thermal units
Travel times of parcels = from system hydraulics

Probabilistic Elements = means and variances of the error terms for the measurement and the systems model*

- Input assumptions should be identified and prioritized as to their potential level of impact

Response: Input assumptions have been identified in the Report.

- Data limitations, assumptions, and approximations inherent in the modeling process introduce errors and inconsistencies into the assessment. Accumulated error can lead to the results of the model being unacceptable or incomplete. Based on that statement, the potential error sources for this analysis should be identified.

Response: Input assumptions have been identified in the Report.

- The input conditions should be identified.

Response: Input conditions have been identified in the Report.

- How were the three flow levels in the Columbia and Snake Rivers chosen? Are they the boundaries of operation? Averages? High, medium and low flows?

Response: The objective of the assessment was to evaluate the relative impact of dams and tributaries on water temperature given existing management of the system and variability in meteorology and hydrology represented by the period 1975-1995.

- Do these flow levels represent specific geomorphic constraints? Specifically is the high flow considered in the flood plain?

Response: Actual flows for the period 1975-1995 were used in the analysis.

- Figures 6 through 13 relate to the simulated and observed water temperatures for the period of 1990-1995 for eight dams. In some instances the simulated results do not match the observed for both high and low periods. Is this difference due to lack of data? Does the model have less ability to accurately predict at the high and low ends of the projection?

Response: Mean and standard deviations of the seasonal differences between observed and simulated have been added to the Report. In general, mean error is lowest during the summer months and of the order of 0.2 to 0.4 °C.

- The concept of the innovation vector analysis and the application to figures 14 through 21 needs to be explained in more detail. Is this application identifying seasonal shifts in temperature? What does the scale represent (-3 to +4)?

Response: A more detailed discussion of the innovations sequence has been added to the Report.

- Figures 22 through 29 are comparisons of actual and simulated innovations. These graphs are hard to read in black and white and perhaps either radically changing the line thickness or using different colors would make them more useful. None-the-less, it appears that the comparison between the observed and simulated is not a good fit. These graphs need to be explained in the result section to help understand their relevance to the evaluation of system model bias and error. Table 11 helps but I really think that the difference between the sample and theoretical variance needs to be explained ir relationship to the modeling effort.

Response: An effort has been made to improve the graphics and t discussion of system model error.

IIX. MODEL APPLICATION

Comments:

- How were these three scenarios developed?

Response: The three scenarios were developed as a controlled experiment to assess the relative importance of dams and tributaries on the thermal energy budget of the Columbia and Snake rivers.

- Is the 16-degree Centigrade temperature regime from the tributaries achievable?

Response: The rationale for choosing the 16-degree Centigrade constraint is based on the State of Washington's water quality criterion for water temperature in Class AA Extraordinary waters. However, the choice of this constraint was not meant to imply that this temperature was achievable. Rather, it was meant to assess what impact the lowering of water temperatures in tributaries would have on water temperatures in the main stems.

- How was the benchmark of 20 degrees Centigrade chosen (page 18)?

Response: A discussion of the rationale for the benchmark of 20 degrees Centigrade has been added to the Report.

- Five areas of issue were identified that require subsequent analysis for future evaluation of Columbia and Snake River temperatures. Can the five areas be prioritized as to their:
    - Level of impact to results
    - Level of impact as related to model calibration
    - Level of statistical importance to evaluating the results

Response: This setting of priorities would be the next of phase of the problem assessment for a TMDL.

- Was irrigation return flow considered important in the analysis?

Response: Irrigation return flow was not considered to be important other than as it affected the aggregate groundwater return flow and temperature.

- Was reservoir retention and operation determined to be an important component of the heat budget

Response: Reservoir retention and operation plays an important role in the temperature regime of the Columbia and Snake rivers. Reservoir geometry and operation determine the travel time through the two systems and also the rate of heat exchange across the air-water interface. The Report concludes that these are major factors leading to alteration of the temperature regimes of the two rivers.

- Was evaporation considered to be an important element in model calibration?

Response: Evaporation rates were adjusted, relative to the Lake Hefner coefficients, to reduce the bias in the difference between observations and simulations. The evaporation coefficients used in the analysis have been added to the Report.

- Figure 30 through 35 and 36 through 41 are really the essential elements of this assessment. I would suggest overlaying the graphs (to show total change) or developing a table for the differences between the five dams and the frequency of exceedance would be useful for the RESULTS section. I also think an arrow indicating the direction of flow (upstream to downstream) would be helpful for interpretation sake.

Response: Comment noted.

IX. RESULTS

Comments:

- Summarize the results with the graphics developed. Specifically it would be useful to overlay figures 30-32, 33-35, 36-38, 39-41. In this way each of the scenarios can be addressed with specific reference to changes predicted.

Response: Comment noted.

- Develop specific headings for each of the three scenarios and identify specific graphics (see above) to assist in evaluating them.

Response: Comment noted.

- A discussion on the model error as related to the results should be developed. Are the results statistically valid?

Response: Discussion of model error has been added to the report. The question of statistical "validity" is more difficult since no protocols have been established for deciding what is a "valid" model. A more thorough discussion of this issue will be part of the problem assessment for any TMDL which might be performed.

- Are the results for levels of exceedance within the statistical ability of the model? Specifically is a 1.4 degree variance at Grand Coulee dam supportable with the level of effort in a screening model? The point is it is that it might not be the actual number that is appropriate but instead be the trend that is seen. With the level of error imbedded in the coefficients and in the model-input data, it might not be safe to say that the actual change is 1.4 degrees. Instead it might be more appropriate to indicate that a thermal increase occurs and exceeds the threshold for specific salmonid species and life stages.

Response: References to the level of effort as a screening model have been removed from the Report. The complexity of the model is similar to that of other analyses (Systems Operation Review, Lower Snake River Temperature and Biological Productivity Modeling) of the Columbia and Snake rivers. The results of the analysis imply that structural differences in the system resulting from dam construction and operation lead to increases in the thermal energy of the two rivers compared to that of the unimpounded rivers. The uncertainty in the state estimates provides a measure of how much improvement in the model and/or the measurement system is necessary to reduce the risk in decision-making.

- Did the models perform as you hoped or was there a need to manipulate the coefficients to allow the model to balance?

Response: The process used in this analysis was to use available measurements to estimate certain model parameters and then to use the resulting modeling to assess the impacts of tributaries and dam construction and operation on water temperatures. The uncertainty analysis was conducted to provide a means for assessing the risks in using the model as decision support tool.

- No discussion is included on how good the model did versus the actual temperatures. This should be a separate section on Model Validation in the result section. The results of the modeling are only as good as the model predictions.

Response: See above discussions on the issue of model "validity". In addition, an extensive discussion of the philosophical problems with assessing model "validity" or "acceptability" has been added to the Report.

- A separate heading on the results from figures 30-35 and a table would be helpful.

Response: Comment noted.

- A separate heading on the results presented in figures 36 - 41 and a table of results should be developed. Specifically in addressing whether the changes what are documented between 36 and 37 are statistically significant.

Response: Comment noted.

- When do the results exceed the 20-degree Centigrade threshold?

Response: Only the frequency and magnitude of temperature excursions are provided in the Report.

- How much natural (pre-project) variability can explain away the thermal increases (without dams) that is predicted?

Response: Pre-project conditions were not evaluated in the Report.

- What figure 39 tells me is this:
  - Water warms as it goes downstream
  - There is a thermal jump at McNary dam and this is due to the Snake River influence
  - There is a thermal jump at McNary without the four dams on the lower Snake River

  WHAT IT DOES NOT TELL ME is how significant the thermal difference is and if the model is good enough to believe.

Response: See previous discussion on model "validity".

X. CONCLUSIONS

<u>Comments:</u>

- The conclusions are supported by the data presented.

Response: Comment noted.

- A DISCUSSION section should be included here to help interpret the results and conclusions drawn.

Response: Some additional discussion and interpretation of results has been added to the Report.

- Questions arise as to the level of detail of the model results as related to the changes identified. For example, is the model sensitive enough to allow for percentages as low as 1-3% to be valid? No results were presented that evaluated the level of change in model results that could be realized with small incremental changes in the model parameters. A section in the conclusions on the Model should be developed. This section would address how good you feel the model is as related to the applications.

Response: The uncertainty analysis aggregates all the uncertainty and variability in the model, including that of uncertainty in the model parameters. Some additional discussion and interpretation of results has been added to the Report

- Is the 1-3% increase due to Snake River dams (conclusion 3) due to upstream Snake River dams?

Response: The effect of the upstream Snake River dams is the same for all three scenarios.

- What is the level of error associated with the results and the therefore the conclusions?

Response: See previous discussion of model "validity" and level of error.

- Are there limits to the use of this model based on the results presented? My fear is that without identifying some limits anyone may think that it is applicable. For protection sake it might be wise to address future uses of the model (i.e. limits, assumptions, etc.)

Response: See previous discussion of model "validity" and level of error.

- Is a conclusion that the reservoirs increase the thermal condition in the river? If so then the reservoirs are indeed heat sinks and even though they may be run-of-the-river they do have an influence on the thermal character of the river. Therefore I strongly urge that you include (as I stated earlier) a section on reservoir dynamics.

onse: The results of the analysis lead to the conclusion that the struction and operation of hydroelectric facilities on the Columbia l Snake rivers results in increases in the thermal energy load of the stem compared to the unimpounded system. An analysis of reservoir ynamics would be an important part in the next phase of a problem assessment for the Columbia and Snake rivers.

- A discussion on the changes that occur at McNary as a result of Snake River inflow would be helpful.

Response: Discussion of the effect of the Snake River on the Columbia River has been added to the Report.

April 19, 1999

Water Management Division

Mr. John Yearsley
Environmental Scientist
U.S. EPA Region 10
1200 Sixth Avenue
Seattle, Washington 98101

Dear Mr. Yearsley:

The U.S. Army Corps of Engineers (Corps) district and division staffs have reviewed the Environmental Protection Agency's (EPA) Columbia River Temperature Assessment Simulation Methods of February 1999 Draft report and offer the following comments. We would have liked to do a more comprehensive review of the Environmental Protection Agency (EPA) model by actually performing test runs. Such a hands-on review was not possible, as we were unable to obtain a copy of the program codes and its input/output. In general, not enough modeling details were provided to allow for a complete review.

Response: As participant in the regional temperature team, EPA Region 10 presented four workshops in Portland on the development and application of the model prior to the release of the draft report. Technical representatives from the USACE attended all four of these meetings. In June 1998, EPA Region 10 technical staff (John Yearsley) installed and successfully executed the Columbia River temperature assessment model on USACE computers in the North Pacific Division offices in Portalnd. USACE technical staff were present during the installation and successful execution and had ample access to the computer code prior to the release of the draft report.

We appreciate EPA's efforts in developing this model, which when fully developed is likely to be widely used in the region. The model certainly has potential for use in developing management strategies for the cold water resource releases at several storage projects in the Columbia River Basin.

We have four areas of concerns:

(1) Geographical coverage,
(2) Extrapolation to pre-dam conditions,
(3) Output accuracy, and
(4) Presentation of model results.

(1) Geographical Coverage. The model encompasses Grand Coulee tailwater to the mouth of the Columbia and the Clearwater confluence at the Snake River down to the confluence of the Snake and Columbia River. We believe this geographical delineation is unduly restrictive in view of EPA's regional responsibilities. Reaches of the upper-Columbia and mid- and upper Snake River contain deep storage (e.g., Grand Coulee, Dworshak, and Brownlee) and tributaries which may contribute to improving river temperature downstream. We suggest including these reaches in your geographic scope.

Response: The objective of this phase of the temperature assessment was to evaluate the effect of dams and tributaries on the the temperature regime of the Columbia River from Grand Coulee Dam (R.M. 596) to Bonneville Dam (R.M. 145) and of the Snake River from Lewiston, Idaho (R.M. 139) to the confluence with Columbia River near Pasco, Washigton (R.M. 0.0). To accommodate the effects of temperature control from Dworshak Dam on the Clearwater River, the mathematical model was extended upstream on the Clearwater River to the its confluence with the North Fork and upstream on the Snake River to its confluence with the Grand Ronde River, just downstream from Hells Canyon Dam (R.M. 168).

(2) Extrapolation to Pre-dam Conditions. There was little information on how the model was used to replicate pre-dam runoff and water temperature conditions. More explanations would enhance user confidence in this type of model application. Explanations should cover theoretical approach, model calibration, data used and accuracy of the validation (e.g., what data were used for the model calibration and how good were the results?) and concrete examples of previous similar model applications.

Response: Pre-dam conditions were not simulated in this analysis. The water temperature of the Columbia and Snake Rivers was simulated for a 21-year period using flows and meteorology from the period 1975-1995. Water temperature was *simulated for scenarios, using the meteorology and hydrology from this 21-year period with (1) the dams in place; (2) dams removed; (3) and dams in place and tributaries constrained to have water temperatures of 16 °C or less.*

(3) Output Accuracy. On the whole, one can only be impressed by the model's capability to replicate the observed daily water temperatures in time and in amplitude. On the other hand, the 5-year time scale used does not provide a closer look at the model's performance in simulating daily data. It would be useful to show the max/min/mean errors achieved for the critical seasons (summer and fall), and the accuracy achieved in replicating benchmark conditions. Finally, a discussion of the nature and extent of the model's limitations (all models have limitations) would be helpful.

Response: An expanded discussion of model limitations has been included and an appendix with statistics for evaluating model performance has been added.

(4) Presentation of Model Results. The information EPA is trying to convey may not be clear enough for the general reader. We suggest using daily water temperature plots on a year-round basis to provide a clearer explanation of changes in the timing and magnitude of pre- and post-impoundment daily water temperatures. Multi-year statistics based on a given set of standards (64 to 66 degrees F) may be too abstract for most readers. It should be clear to the reader that the temperature standards were not attained under pre-dam conditions. Based on the information presented in the current draft, the general reader could conclude that the dams have significantly raised water temperatures, which is not necessarily correct.

Response: Pre- and post-impoundment conditions are not evaluated in this report. The objective of the report is to assess the relative impacts of the impoundments and tributaries on the thermal regime of the Columbia and Snake rivers. The report shows that the benchmark of 20 °C is exceeded under all conditions at some locations.

Finally, because there are water temperature models already in use, we suggest EPA discuss in more detail the purposes this particular model serves. It is important that all models convey a consistent message about water quality changes.

Response: The Report discusses other temperature models and provides a rationale for the choice of the method used in the Report. This model is being used to assess water temperature in the Columbia and Snake rivers in support of Section 303 (d) of the Clean Water Act and it is EPA's policy to perform peer review on technical products which may be of major importance. So far as the author knows, none of the other models used for policy-making in the Columbia and Snake rivers, has been subject to formal peer review.

More detailed comments are provided in the attachment.

Thank you for the opportunity to comment on this modeling effort. If there are any questions, please contact Bolyvong Tanovan at (503) 808-3938.

Sincerely,

James E. Crews, P.E.
Director, Engineering and
Technical Services

Enclosure

ENCLOSURE

SUBJECT: Comments on the EPA's Columbia River Temperature Assessment Simulation Methods (February 1999 Draft)

1. Page 1. Introduction. The draft refers to "*... the Snake River from Lewiston... to its confluence with the Columbia River as water quality limited for water temperature under Section 303(d) of the Clean Water Act,*" quoting an analysis made by Washington DOE. Even though the Snake River is not listed on Idaho's 303(d) list as water quality limited for water temperature, water temperature conditions on the upper Snake River Idaho are not that much better. By not covering the upper Snake River, the report understates the extent of the water temperature problem.

Response: See response to <u>Geographical Extent</u> given above.

2. Sources contributing to changes (2), suggest editing as follows: "*modifications to the natural river system to generate electricity, provide irrigation water for farmlands and facilitate navigation.*"

Response: The text was modified as suggested.

3. Page 2, first line. "*...changes in the water temperature regime of the main stem Columbia River in Washington and Oregon and in the Snake River in Washington.*" This geographical delineation is unduly restrictive (see above). EPA has regional responsibilities extending to Idaho. Excluding Idaho river reaches, which also contain storage reservoirs and tributaries that may be part of the solution, does not seem logical.

Response: See response to <u>Geographical Extent</u> given above.

4. Page 3, 2nd paragraph. "*Coastal stations are typically higher*. To be more accurate, shouldn't this be revised to read "*Precipitation recorded at coastal stations is typically higher.*"?

Response: The text was modified as suggested.

5. Page 3, 4th paragraph. "*This is particularly true of the Willamette River, which occasionally reaches flood stages even with flood control available from system reservoirs.*" This sentence could be misleading. The Willamette River is a system in itself that depends little (except near its confluence with the Columbia River) on other system reservoirs for flood relief. The Willamette River can reach flood stage when storage projects are releasing minimum outflow and storing at their maximum rate. Suggest revising.

Response: The text was modified as suggested.

6. Table 1. Suggest adding "(cfs)" to the Average Flow column's title.

Response: The text was modified as suggested.

7. Table 2. Suggest you refer to Dworshak and Brownlee Reservoirs in the footnote as potential contributors to alleviating water temperature problems in the Columbia River Basin.

Response: The objective of this study was to assess the impacts of dams and tributaries on the temperature regime of the Columbia and Snake rivers. Strategies for attaining water quality standards would be considered in an analysis of options available for temperature control or modification.

8. Page 3, 6th paragraph. "*The projects develop approximately 1,240 feet of the 1,290 feet of hydraulic head.*" Suggest adding qualifiers to make this sentence easier to understand.

Response: The text was modified as suggested.

9. Page 3, 7th paragraph. "The system has the capacity...."

Response: The text was modified as suggested

10. Page 4, 2nd paragraph. "The nature of water quality problem is described in the list... prepared by the State of Washington's Department of Ecology." Suggest stating that this is only one source of information. Other states in the Columbia River Basin have issued similar lists.

Response: Reference of the Oregon DEQ 303 (d) list of impaired waters was added.

11. Page 4 and Table 3. Need to clarify if the information presented is based on WA 1996 approved or WA 1998 candidate 303(d) list.

Response: Clarification was provided as suggested.

12. Page 4. "*Changes in temperature and gas pressure of water released from hydroelectric projects have an impact on the aquatic ecosystem of the Columbia River system, particularly salmon and steelhead.*" Suggest that a reference for this statement be included. The impact is probably felt by all aquatic biota in vicinity, "including salmon and steelhead."

Response: The discussion of temperature impacts on salmonids was expanded and additional references provided.

13. Page 5, 2nd and 3rd paragraphs. We question the relevance of the reference to the Barnwell and Krenkel (1982) work. These authors have not pioneered water quality modeling and

application to management, nor have they established any particular yardsticks that would influence current day's mathematical modeling. Therefore, when it is stated that the "*study objectives require a level of spatial and temporal complexity... greater than for the screening models described by Barnwell and Krenkel,*" this invites more questions than answers.

Response: References to "screening models" and the specific reference to Barnwell and Krenkel were removed as suggested.

14. Pages 5 and 6, System Boundaries. On page 5, it is stated that the "*boundaries of the Columbia River system ... include the Columbia River from the International Border to.*" Yet, on page 6, it is also stated that *"the system boundaries for the model of the run-of-the-river segments are from the tailwaters of Grand Coulee to Bonneville.* "Suggest additional clarification/explanation as to why different starting points were used in the study and in the model.

Response: The text was modified as suggested

15. Page 6 and Table 4. What criteria were used to exclude a tributary from the thermal analysis? For instance, Walla Walla River was excluded from Table 4, but Tucannon River was included.

Response: The criteria for including tributaries were (1) size of tributary and (2) availability of data. In general, availability of data was the controlling criterion. The Walla Walla River and groundwater return flow have been added to the report.

16. Pages 7 through 12. We have not checked the accuracy of your theoretical approach and assume that other reviewers will provide appropriate comments on that approach.

17. Page 9, 1st paragraph. A little more explanation of the Karman (*sic*) filter "recipe" and why it is applicable to the Columbia-Snake mainstem would be helpful. Which other comparable river basins has this filter "recipe" been applied to?

Response: The reference to the Kalman filter as a "recipe" has been removed. The Kalman filter is applicable to any problem that can be formulated in terms of the state-space model. Selected references to the use of the Kalman filter in water quality applications have been included.

18. Page 13, Meteorology. Would be interesting to indicate how many hydro-met stations were used and how those stations were assigned to individual basins. Also, a discussion of historical stream temperature would be relevant. For instance, did the model use any data at the mouth of the Snake River before the lower river projects were built?

Response: A description of meteorological stations used in the temperature assessment is in the report. The objective of this study was to assess the impact of dams and tributaries on water *temperature of the Columbia and Snake rivers in a controlled experiment. Davidson (1964) provides the most comprehensive analysis of available historical water temperature data.*

19. Page 16. Water temperature probes used in the Corps total dissolved monitoring program are generally placed under 15 feet of water (not 10 feet).

Response: The text was modified as suggested

20. Page 18, paragraph 4. Indicate the frequency that tributary temperature exceeds 16 degrees C.

Response: The benchmark of 16 °C used for assessing the impact of tributaries was based on the water quality criterion for Class AA waters, as defined in the State of Washington's water quality standards (Chapter 173-201A-030 (2) (b) of the Washington Administrative Code (WAC)).

21. Page 18, paragraph 5. What was the selected 20-degree C benchmark based on?

Response: A discussion of the rationale for selecting the20 °C was added.

22. Page 19. Water Balance. "*Withdrawals for irrigation, groundwater return flow and miscellaneous tributary flow ... comprise an estimated 5-7% of the flow increment to the Columbia River.* "Would be interesting to provide some ideas about the accuracy of the water accounting achieved with the model from station to station and indicate how "negative" local inflows were treated.

Response: The accuracy of the water balance is of the order of 5-7%.

23. Page 20, 2nd paragraph. "*increased response time associated with shallower depths and higher velocities.*" This seems counter-intuitive if talking about response time of temperature changes. Maybe, the word "*response*" needs to be better defined.

Response: The text was modified in an effort to clarify this.

24. Page 20, third paragraph indicates that the duration of exceedance with the dams in-place on the Lower Snake is greater than without dams. Recent LSRFS model (WQRRS) simulation results for free flowing condition also show less days above 20°C than the empirical data from the same years for impounded condition, but the difference is not as substantial. This may be due to the fact that we are working with years that all include augmentation releases from Dworshak in the late summer (1994 and later). This would lead to "artificially" fewer days of exceedance of 20°C than before the inception of operation of Dworshak in that manner.

Response: The referenced LSRFS simulation results have not been subject to a formal peer review process and the results, therefore, can be considered only as preliminary. In addition, studies reported in Appendix D of this report suggest that the WQRRS model may give questionable results for problems that include inputs with time scales of less than 10-20 days.

25. Also in that third paragraph of the Results is a finding that the magnitude of exceedance is greater with impoundments than without dams. Some historical data do not seem to support this generic conclusion. Temperature measurements taken at the mouth of the Snake River in the late 1950's were as high as 79.5°F, and there was almost always a summer peak higher than any we see today. Suggest additional clarification on this point as it makes intuitive sense, that the temperatures wouldn't necessarily get as high with impoundments. The temperature can't follow ambient conditions as closely because of the buffer provided by the lower surface area to volume ratio of an impoundment. An impoundment will cause increased retention time and some warming, and certainly delays cooling.

26. Page 19, paragraph 5, last sentence. "*Groundwater temperatures are likely to be lower than the main stem in the summer and higher during the winter.*"

Response: The text was modified as suggested.

27. Figures 6 through 13. On the whole, one can only be impressed by the model's capability to replicate the observed daily water temperatures in time and in amplitude. On the other hand, the 5-year time scale used does not provide a closer look at the model's performance in simulating daily data. Would also be useful to show the max/min/mean errors achieved for the critical seasons (summer and fall).

Response: The text was modified as suggested.

28. Information on how pre-dam water temperature and flow conditions were modeled was very scarce. What data were used for the model calibration and how good were the results? We were particularly concerned about the rather limited impacts shown by the model as a result of Dworshak operation. To gain a straightforward comparison of the effects of the dams and of Dworshak augmentation on Snake River water temperature, we suggest that three scenarios be simulated: (a) river without dams, (b) river with dams but prior to Dworshak augmentation, and (c) river with dams and with Dworshak augmentation.

Response: Pre-dam conditions were not simulated in this temperature assessment, as described previously. Effects of strategies which attempt to modify or control temperatures in the Columbia and Snake rivers will be evaluated in further studies.

29. Figures 14 through 29. What is the practical significance of the "*innovations variance*"?

Response: The discussion of the practical significance of the innovations sequence and its use was added.

30. Figures 30 through 41. Same comments on the frequency plots. What simulation years were included in the statistics (the same 1990-1995 as in the other figures?). Also, what geographical location does the word "Confluence" shown on the X-axis represents? Snake/Clearwater confluence? Assuming that was indeed the case, Dworshak reservoir operation does not seem to have any impact at all --which deserves some explaining.

Response: The text was modified as suggested and the geographical scope of the model was expanded to include the operation of Dworshak reservoir.

31. It would be interesting to once more explore with the model how the lower river might be cooler if upper river storage projects had a selective withdrawal system.

Response: Comment noted

REVIEW OF COLUMBIA RIVER TEMPERATURE ASSESSMENT: SIMULATION METHODS

APRIL 2000

Prepared for:

Potlatch Corporation
Lewiston, Idaho

Prepared by:

Peter Shanahan, Ph.D, P.E.
M. Bruce Beck, Ph.D.

HydroAnalysis, Inc.

481 Great Road, Suite 3
Acton, Massachusetts 01720
(978) 263-1092
fax: (978) 263-8910
e-mail: shanahan@ma.ultranet.com

REVIEW OF
COLUMBIA RIVER TEMPERATURE ASSESSMENT: SIMULATION METHODS

INTRODUCTION

This document provides a review of the report "Columbia River Temperature Assessment: Simulation Methods" prepared by John Yearsley of U.S. Environmental Protection Agency Region 10 in February 1999. The report employs sophisticated modeling and parameter estimation techniques to assess temperature in the Columbia River system and place bounds on the uncertainty in predicting temperature.

The introduction to the EPA report lists the following three sources that may contribute to changes in the temperature of the Columbia and Snake Rivers: impoundments, hydrologic modifications, and watershed modifications. The report does not consider a potentially significant contributor to temperature change, global climate change. The report indicates that it is the objective of the study "to assess the relative importance of these sources with respect to changes in the temperature regime of the main stem Columbia River in Washington and Oregon and in the Snake River in Washington." This objective is partially accomplished. The report limits its focus to the effects of dams and a blanket consideration of contributing tributary temperature.

The number of simulations in the study and their sophistication are insufficient to achieve fully the stated objectives and the study should be viewed as a screening level analysis to identify potential factors that affect river temperature. The report notes on page 5 that the model is indeed a screening model, designed only "to identify critical areas for additional analysis." Thus, the model results should be recognized to be approximate and exploratory rather than definitive. This report would be an inappropriate basis for policy decisions other than the identification of areas for further research.

MATHEMATICAL MODEL DEVELOPMENT

The author elected to write a new computer program for modeling temperature rather than relying upon an established model such as QUAL-2E (Brown and Barnwell, 1987). As explained by author, his approach provides some legitimate advantages for the problem in question. Nonetheless, the failure to use an established model that has experienced substantial use and prior review places an added burden on the author to verify his model.

Typically, computer-model verification is accomplished by using the model to simulate one or more problems for which there is a well-known analytical solution. Close agreement between the model solution and the analytical solution "verifies" the model; that is, it demonstrates that the model accurately solves the type of problem that it was designed to solve (National Research Council, 1989, pp.235-236). The greater the number of such test cases performed, the more confident one can be of the model's accuracy and validity. In essence, verification is a process to guarantee against erroneous computer code. It protects against the situation in which erroneous code begets erroneous results, known in computer science by the shorthand, "garbage in, garbage out."

No verification of the Columbia River temperature model was provided in the reviewed document. Use of the model for screening purposes does not obviate the need for verification. Without such verification, the model's accuracy cannot be assessed and the model cannot be relied upon. The U.S. EPA should be requested to provide documentation of model verification in accordance with accepted standards.

Response: A number of tests of the computer code were performed. In addition, the frequency response of the numerical method was evaluated and compared to the frequency response of other models that have been used for policy-making in the Columbia and Snake rivers. The results of these tests have been added to the Report.

While potential programming errors could compromise the model's validity, the model has a solid technical foundation. The model equations as presented in equations 1, 2, 4, 13-15, and 17-21 are well known and accepted. The equations for stream temperature are solved in the EPA model using a mixed Eulerian-Lagrangian solution scheme. This approach is much less common in surface-water-quality modeling than an Eulerian approach, but is nonetheless entirely valid. The Eulerian approach, which is used for example in QUAL2E, solves for temperature at fixed points along the river. In essence, the Eulerian solution looks at the entire river at a single time. In the Lagrangian approach, the numerical solution follows a particular parcel of water down the river. In essence, this technique looks at this single parcel, following it in time and space as it moves downriver with the streamflow. Travel time for the parcel, τ, and its location along the river, x, are directly related by the stream velocity, U: $x = U\tau$. The Lagrangian approach takes advantage of this relationship to eliminate the distance variable from the equations being solved.

The mixed solution in the Columbia River model uses a Lagrangian alg model the effect of water flowing downstream in the river, but an Eulerian algo model mixing. The main advantages of the mixed Lagrangian-Eulerian techn general, are accurate representation of dispersion and the ability to model sharp trar in concentration. These are not factors in the current model version since diffusion is not modeled and the concentration (temperature) changes gradually. Thus, although these factors are cited on page 10 of the report, the real advantage of using this solution in the Columbia River model seems to be that it simplifies the accompanying Kalman filtering analysis of model parameters. The role of Kalman filtering in the Columbia River model is further described below.

DATA SOURCES AND PARAMETER ESTIMATION

A multi-step process that considered the statistical properties of temperature prediction and measurement and their uncertainty was used to establish the various parameters in the model. First, model parameters were set based directly on the available data and literature. This first step is called the deterministic parameter estimation. Next, these parameters were adjusted to produce results deemed "unbiased." In essence, this is the same as the manual calibration process used in traditional water-quality modeling. Finally, the variance of the systems model (i.e., the model for river temperature) was estimated as the third and final step of the parameter estimation process. Variance is a statistical measure of the variation of the model results around its mean prediction and depends upon the error in both the model and field data.

The Columbia River model requires a large number of input data to represent the river's geometry and the basin hydrology and meteorology. As discussed on page 13 of the EPA report, numerous past studies of this river system provide an unusually rich database for modeling. The first step in the model development, deterministic parameter estimation, was thus a fairly straightforward and familiar process. A few exceptions are discussed below.

On page 14, the report describes the basis for estimating the river's hydraulic parameters as a function of streamflow under the dam removal scenario. These estimates were made for flows of 60,000, 120,000 and 240,000 cfs in the Snake River. Unfortunately, the minimum flow of 60,000 cfs exceeds the average monthly flow in the Snake River below Ice Harbor Dam for the months of July through March and is more than twice the monthly average flows for August, September, and October. The model hydraulic parameters thus may be inappropriate for modeling the dam removal scenario during the critical low-flow months. This is significant because these months are also the most likely to experience temperature exceedances.

Response: The hydraulic parameters for the unimpounded Snake River, estimated for the flows of 60,000, 120,000 and 240,000 cfs were used to estimate the hydraulic conditions for 10,000 cfs. These estimates were compared with the results of simulating steady flow with HEC-RAS at flows in the Snake River of 10,000 cfs. A comparison of the two methods showed that the hydraulic parameters used in the report gave results similar to the simulated results from HEC-RAS.

In computing surface heat transfer, the author has appropriately opted to use a heat budget model when simpler, but less accurate, approaches are sometimes used. The heat budget model, which is presented on page 15 and in equations 17 through 21, is based on an authoritative reference (Wunderlich and Gras, 1967) although a somewhat later and more accessible version of the same reference exists (TVA, 1972). The presentation in the EPA report is incomplete, however, in that the report fails to define all variables and units in the equations. This creates some uncertainty as to exactly what relations were actually used. One critical factor in calculating surface heat transfer involves the evaporative heat flux in equation 19. There are multiple formulas for the dependence of this flux on wind speed and the formula chosen may significantly alter the computed heat transfer (Shanahan, 1985). The EPA report fails to clarify the exact formulation used. With respect to the shortwave radiation in equation 17, the atmospheric attenuation coefficient has been shown in some studies in the Pacific Northwest to take on unusual values owing to haze (Findikakis *et al.*, 1980). This may be a factor in some areas of the watershed, which would necessitate special parameter selection.

Response: Peer reviewed methods for determining the various components of the heat budget were used in this report. Additional discussion of parameters, including definition of variables and their units, was added.

On page 15, the report states that **daily** temperature values are not always available for the upstream stations that form the model boundaries. The report then presents, on page 16, an empirical equation for **weekly** temperature that was used to fill in missing temperature values. The mismatch between the daily and weekly periods is significant in that there may be a significant time lag between meteorological conditions and the resulting stream temperatures. As a result, the relation for weekly conditions is likely to be substantially different from the relation for daily conditions. It is unclear from the report whether daily temperature values were actually derived or whether weekly temperatures were used during periods of missing data. In either case, there is substantial margin for error in fixing the temperature at the model inflow points.

Response: A peer reviewed method was used to estimate weekly water temperatures for tributaries with missing data. The report acknowledges the uncertainty associated with these estimates.

Under the subheading "Systems Model Bias and Error" starting on page 17, the report describes how Kalman filtering was used. Kalman filtering augments the supposedly certain prediction made by a traditional water-quality model with a probabilistic prediction that recognizes multiple sources of uncertainty. In a traditional model, the temperature at some time step k is determined from the temperature at the preceding time step $k-1$, using the model equations and known parameters. In fact, however, the temperature at time step $k-1$ is known imperfectly, the parameters for stepping from time step $k-1$ to time step k are uncertain, and even the model formulation itself probably has errors. If the model results are compared with field measurements, the field measurements must also be recognized as having some error. The Kalman filtering approach recognizes these various sources of error and incorporates them into the model formulation. The result is an estimate of the uncertainty in the model predictions that can be used to help guide the calibration procedure.

Kalman filtering is a complicated and specialized technique. Accordingly, expert review of the EPA's application of Kalman filtering in the Columbia River temperature analysis was sought from Professor M. Bruce Beck of the University of Georgia. Dr. Beck is an internationally-known specialist in the application of Kalman filtering to surface-water quality. Dr. Beck's review is appended to this review. Dr. Beck finds no fault in the technical aspects of the Kalman filter analysis, but raises some cautions as to the interpretation of the results. These cautions are pointed out in the discussion that follows.

The EPA study used Kalman filtering in an approach that closely follows that presented by Van Geer *et al.* (1991) for a ground-water modeling application. Despite the change in environmental medium, the approach remains valid for the EPA application. As presented by Van Geer *et al.*, Kalman filtering provides information on the uncertainty of the model prediction and helps guide the calibration process. According to Van Geer *et al.*, one can achieve a better calibration using Kalman filtering than the traditional, deterministic approach. In a personal communication, Dr. Beck has indicated his strong disagreement with this assertion.

Response: The approach used in the report was based on a number of peer reviewed studies, including that of Van Geer et al (1991).

The Kalman filtering procedure is complicated, as is implied in equations 5 through 12 on page 9, and the description in the EPA report is spare and difficult to follow. Accordingly, the procedure is summarized in the next three paragraphs, which are based on the more lucid description in the paper by Van Geer *et al.* (1991). These three paragraphs, which are quite technical, can be skipped without loosing the overall sense of this review. With respect to the procedure described in the EPA report, equations 8 and 12 include errors. In equation 8, the second instance of f_{k-1} should instead be its transpose, f_{k-1}^T. In equation 12, $\underline{v}_k$ and $\underline{z}_k$ are vectors and should both be underscored.

The Kalman filtering process, as described in equations 5 through 12, marches through time in discrete time steps. It consists of two sub-steps at each time step: first a prediction is made strictly from the model equations, and second it is corrected based on the measured data.

The first sub-step is the prediction. At each time step k in the temperature simulation, the temperature at the various measurement locations along the river (represented by the vector $\underline{T}_k$) is predicted with equation 7 as a function of the system matrix (f_{k-1}) and the temperature at the last time step, $\underline{T}_{k-1}$. The system matrix is simply the temperature equations 2 and 4 in another form. In parallel with the temperature prediction at time step *k*, the uncertainty in the prediction is estimated with equation 8. The uncertainty in the temperature is a matrix, P_k, in which the diagonal elements are the variances of each temperature value (i.e., the temperature at each station) and the off-diagonal elements are the covariances between the temperatures at different stations. This matrix is known as the error covariance. Like the temperature vector, the error covariance matrix is predicted based on its value at the last time step.

Following the predictor sub-step is the corrector or update sub-step. Here, the predicted temperature is updated with the actual temperature measurements, $\underline{z}_k$, using equation 9. Equation 9 is simply a weighted average of the predicted temperature and the measured temperature, but with the weighting changing as the simulation progresses. The weighting is captured in the so-called Kalman gain, K_k, which is also a matrix and is calculated in equation 11. In the actual computational sequence, equation 11 is completed before equations 9 and 10. The error covariance is similarly updated in equation 10. At the end of the update sub-step, the calculation for time step *k* is completed and the process begins again with the predictor sub-step for time step *k+1*.

An outcome from Kalman filtering is the innovations sequence, equation 12, which shows the error between the measured and predicted temperature at each location along the river at each time step *k* in the simulation. The goal in calibrating the deterministic temperature model is to adjust the model so as to minimize the mean of the values of this error term over time. In addition, the error covariance matrix (the Σ_Q term in equations 5 and 8) is adjusted until the innovations sequence satisfies certain statistical properties discussed below. Varying, one at a time, the deterministic model and the properties of the stochastic model error adjusts the model. As described in the EPA report, the only deterministic parameter varied was the meteorological data station assigned to each reach of the river. The assignment of stations was varied manually until, according to the report, "the mean of the innovations vector was small." No specific description of "small" is given, although Figures 6 through 13 allow a visual evaluation of the error. Dr. Beck cautions that Figures 6 through 13 appear to compare updated temperature predictions, and thus may present a more favorable comparison to the field data than is appropriate. As indicated in Dr. Beck's review, the exact character of the simulated values in Figures 6 through 13 should be clarified. 7

Response: Additional discussion of the innovations sequence and the filtering approach was added to clarify these issues.

The stochastic error term that was varied is the estimate of the error in the system model. This error is represented by $\underline{w}_{k-1}$ in equation 5, and it is assumed to be a Gaussian distribution with zero mean and variance Σ_Q. This error is not known at the start of the modeling exercise, so it is given an initial guess and then corrected by trial-and-error based on the results of the Kalman filtering. The corrections entail changing the values of the assumed statistical variance matrix, Σ_Q. The stochastic part of the model is determined to be calibrated when the values of Σ_Q cause the model error computed from the innovations sequence to match the theoretical error predicted by the model. Mechanically, this match is computed using equations 23 and 24 on page 17. As Dr. Beck points out in his review, the values of Σ_Q are expected to differ between simulations of the existing situation with dams and predictions for a future situation without dams. However, it appears that the same values of Σ_Q were used for both scenarios. Dr. Beck also points out that assumptions made concerning the character of covariance terms in Σ_Q are inadequately discussed in the report.

Response: The report acknowledges the uncertainty associated with assuming the systems model error variance, Σ_Q, is the same for all scenarios. This assumption is a result of the absence of water temperature data for scenarios with dams not in place. However, this assumption does not necessarily mean that the variance of the state estimates is the same for all scenarios. The systems dynamics play an important role in the propagation of uncertainty. The systems dynamics for the scenarios with dams in place differ from the dynamics with dams removed and will, in general, give different results for the variance of the state estimates.

The model results raise some questions. The text indicates that data were available for the period 1975 through 1995, but calibration results for only 1990 through 1995 are shown. It cannot be determined from the report whether the entire period of record was used to calibrate the model or if only the 1990-1995 subset of the record was used. It would not be inappropriate to base the calibration on the 1990-1995 period only, since page 16 indicates the data are more reliable then, but the data selection should be clarified.

Response: In the revised report, data collected in association with the total dissolved gas monitoring program for the period 1990-1994 was used in the parameter estimation process. These data were chosen because of their completeness and reliability.

The innovations sequence is a measure of the difference between the temperature predicted by the model and that actually measured. The innovations sequence is shown in Figures 14 through 21 over the calibration period at a number of measurement stations along the rivers. The error is relatively large— greater than 3 or 4 degrees. Moreover, the figures plot a 30-day moving average, implying that some daily values are even more in error. The report is deficient in explaining the meaning, significance, and limitations of these results. The figures illustrate the calibration of the deterministic model where the goal is to get the mean of all plotted values to equal zero. This can be equivalently thought of as getting the area under the plotted curve above the x-axis (0-degree line) equal to the corresponding area below the x-axis. It appears that at some of the stations, the calibration fell well short of this goal. The peaks and valleys in Figures 14 through 21 indicate that the model appears consistently to predict temperatures that are too low in fall, winter, and spring, but too high in summer. Dr. Beck further discusses the lack of coherence between Figures 14 through 21 and Figures 6 through 13 and its implications insofar as relying upon the model predictions.

Response: A number of statistics relating to model performance have been added to the report.

The report is similarly deficient in explaining the meaning, significance, and limitations of the results in Figures 22-29. In essence, these figures report on the calibration of the stochastic model, plotting the results of equation 23 against those of equation 24. A goal of model calibration is to get these to match. Unfortunately, the key of these figures is insufficiently clear to distinguish which plotted line represents which result. The terminology of the figures deviates from that of the text, further confusing the results. As with the results in Figures 14-21, the results in Figures 22-29 show significant variations over time and, in at least some cases, a consistent mismatch.

Based on the comparisons in the figures, it is difficult to assess the quality of the calibration. More information on alternative calibration attempts would be helpful in this regard and would also give a sense of the model sensitivity. As well, segregation of the model-data comparison by month would help in identifying the accuracy with which the highest temperatures are predicted. Model predictions are particularly critical in this range because it is only this portion of the model results that are actually evaluated.

MODEL APPLICATION

On page 18, the report states goals that are not entirely congruent with the objectives stated on page 1. Also incongruent are the conclusions on page 20. While page 5 states that the model is a screening tool capable of identifying areas for further study, the report make no recommendations for further study. Instead, the report lists seemingly firm conclusions— an outcome that is inconsistent with the power and purpose of a screening model.

Model results are shown in terms of the frequency with which a temperature of 20°C is exceeded at the various stations along the rivers (in Figures 30-35) as well as the degree to which the temperatures are exceeded (in Figures 36-41). Simulation scenarios consider the current situation, the situation if existing dams were to be removed, and the situation if temperatures from tributary streams were kept less than or equal to 16°C. The simulations show that the frequency and magnitude with which 20°C is exceeded is decreased by removing dams (other than at the Snake River confluence and Grand Coulee Dam) and relatively unaltered by controlling tributary temperature.

There is a significant mismatch between the way the model was used and the way it was developed that calls into question its predictions. In its use, the model is applied only to evaluate extreme high temperatures that occur in the summer. But, the model's calibration and statistical evaluation were judged in terms of year-round agreement. The statistical measures used in the Kalman filtering evaluate the degree of agreement over the entire year and, for the deterministic model, via summations over the entire calibration period. Thus the summertime predictions, which tend to be high, are offset by the non-summer predictions, which tend to be low. The results presented in the report, however, show extreme temperature exceedances that occur only in the summertime period. Before the model can be confidently used to evaluate temperature extremes, it must be calibrated and checked specifically against periods of high temperature. Dr. Beck confirms this conclusion in his review. ?

This fundamental limitation notwithstanding, the model results predict temperature exceedances (in Figures 36 through 41) that are comparable to the calibration errors depicted in Figures 6 through 21. The "error bars" shown in Figures 36 through 41 may be confusing in this regard. They show the variation of the predicted exceedances around the mean and do not relate to the model uncertainty. However, it is clear from inspection of Figures 6 through 13 that the temperature model makes its poorest predictions at the extremes, yet it is precisely at the extremes where the model is being used. ?

SUMMARY

The EPA Columbia River temperature model uses unusual and technically sophisticated techniques to evaluate the effects of dams and other factors on temperature in the Columbia and Snake Rivers. Because an established model was not used, the Columbia River model should be verified in accordance with accepted practices for model quality assurance and quality control.

Calibration information provided for the model appears to show that the model predicts summertime temperatures that are generally higher than those observed and non-summer temperatures that tend to be lower than observed. However, the model calibration was evaluated in terms of year-round agreement, such that these two systematic errors balance each other. In contrast to the calibration evaluation, the model was used in a predictive mode only to evaluate extreme warm temperatures in the summertime. If the model is to be used primarily or solely to evaluate high temperature extremes, its predictive capability should be evaluated specifically for high temperature.

Errors in the model during summer appear to be comparable to the degree of exceedance predicted for summertime temperature excursion above the 20°C temperature threshold. This relative similarity of model error to the predicted excursion, as well as the mismatch between the calibration focus and prediction focus, indicate that the model results should be considered qualitative at best. As indicated in the report itself, the model is intended as a screening tool to identify areas for further research. As such, the model is not an adequate basis for policy decisions.

In a separate appended review, Dr. M. Bruce Beck focuses on the application of Kalman filtering in the EPA study. Dr. Beck concludes the Kalman filtering is implemented in a technically sound manner overall, but that certain aspects of the application require clarification. He also questions a number of explicit and implicit assumptions regarding the character of error and uncertainty and suggests additional analysis to explore their consequences.

REFERENCES

Brown, L.C., and T.O. Barnwell, 1987. The Enhanced Stream Water Quality Models QUAL2E and QUAL2E-UNCAS: Documentation and User Manual. Report No. EPA/600/3-87/007. U.S. Environmental Protection Agency, Athens, Georgia. May 1987.

Findikakis, A., F.A. Locher, and P.J. Ryan, 1980. Temperature and Turbidity Simulation in Spada Lake. In: H.G. Stefan, editor. *Proceedings of the Symposium on Surface Water Impoundments. June 2-5, 1980, Minneapolis, Minnesota.* Vol. I, Pg. 594-603. American Society of Civil Engineers, New York.

National Research Council, 1989. *Ground Water Models: Scientific and Regulatory Applications*. National Academy Press, Washington, D.C.

Shanahan, P., 1985. Water Temperature Modeling: A Practical Guide. in T.O. Barnwell, Jr., editor. Proceedings of the Stormwater and Water Quality Model Users Group Meeting, April 12-13, 1984. Report EPA-600/9-85-003. U.S. Environmental Protection Agency, Athens, Georgia.

TVA, 1972. Heat and Mass Transfer Between a Water Surface and the Atmosphere. Water Resources Research Laboratory No. 14. Tennessee Valley Authority, Division of Water Control Planning, Engineering Laboratory, Norris, Tennessee. April 1972.

Van Geer, F.C., C.B.M. Te Stroet, and Y. Zhou, 1991. Using Kalman Filtering to Improve and Quantify the Uncertainty of Numerical Groundwater Simulations, 1. The Role of System Noise and Its Calibration. *Water Resources Research*, Vol. 27, No. 8, Pgs. 1987-1994. August 1991.

Wunderlich, W.O., and R. Gras, 1967. Heat and Mass Transfer Between a Water Surface and the Atmosphere. Tennessee Valley Authority, Division of Water Control Planning, Engineering Laboratory, Norris, Tennessee.

APPENDIX

Review of Kalman filtering
by Dr. M. Bruce Beck

REVIEW OF
COLUMBIA RIVER TEMPERATURE ASSESSMENT: KALMAN FILTERING

INTRODUCTION

The purposes of this supplementary review are to:

(i) determine whether the Kalman filter has been used in a correct manner, i.e., in a manner consistent with published (and peer-reviewed) practice;

(ii) assess whether this filtering framework is appropriate for the given task;

(iii) consider alternatives to this framework suitable for any further such studies;

(iv) indicate what might be the desirable subjects of those future studies;

Before reporting on these matters, it will be helpful to review what has actually been achieved through applying the Kalman filter in this context.

GOAL OF EPA REPORT

Without seeking to diminish the significance of, or distract attention away from, variations in stream temperature over the entire annual cycle, let me suggest the real issue here is that of forecasting the maximum temperature over this cycle. Furthermore, if one is conservative in outlook, it may be better to over-predict than to under-predict this maximum. Re-stating this goal is important, because it has a bearing on some of the detail surrounding the way in which the filter has been used in the EPA Report.

Response: The primary objective of the report, as described in the final draft, is to determine, for a given sequence of hydrology and meteorological conditions, the relative impacts of the operation of dams and reservoirs on the thermal energy budget of the main stem Columbia and Snake rivers compared to the impact of thermal input from surface and groundwater sources. Neither forecasting nor *policy-making are objectives of this phase of the study of the thermal regime of the Columbia and Snake rivers.*

"VALUE ADDED" THROUGH USE OF THE KALMAN FILTER

Besides the obviously highly charged political context of this problem, removing dams from the Columbia and Snake Rivers is a rather dramatic piece of policy. It would therefore seem important for all concerned to be reassured that such action is "right for the situation" and to be aware of the risks of "getting it wrong". Considerations of uncertainty and risk, then, are entirely appropriate in such a problem setting. Indeed, to have undertaken this exercise in the absence of such considerations, i.e., under the assumption of an entirely deterministic model, would itself have been an act of engaging in risk-taking (for the EPA). Use of the Kalman filter to address these issues of uncertainty is not usual, but by no means unknown.

Response: The report makes no recommendations regarding the removal of dams from the Columbia and Snake rivers.

The following is the essential role played by the filter in this study, to paraphrase (in perhaps colloquial terms):

> The world is uncertain. We know too that all models are approximations. All sources of uncertainty (approximations, omissions, errors) in the model will be subsumed under the label of the system noise vector ($\underline{w}$). Besides estimates of the model's conventional (deterministic) parameters, we shall therefore need estimates of the variance-covariance properties of $\underline{w}$ (the matrix Σ_Q) in order to account for the manner in which the inevitable residual uncertainty attaching to the model -- even when calibrated -- is propagated forward into forecasts of future behavior (under changed conditions).

Response: With the exception of the usage of the term "calibrated", this statement captures the essence of the method used to propagate uncertainty in the report.

In fact, looking at the source reference of van Geer *et al.* (1991), one might go so far as to say the primary purpose of calibration in the present study is to adjust the estimates of the variance-covariance matrix of the system noise (Σ_Q), with a view to assessing its impact on the uncertainty of the forecasts.

To be clear about what is subsumed under this matrix, we have the following generic sources of uncertainty:

(i) uncertainty in the (deterministic) parameters of the model;

(ii) uncertainty in the measured input disturbances of the model, i.e., here, principally the variations in the temperature of the tributaries;

(iii) uncertainty in all other unmeasured disturbances of the factors affecting temperature (the state variable).

In addition, account must be taken of uncertainty in the system's (past) observations, as must the uncertainty in the initial state of the system, i.e., the values of the spatial distribution of temperatures at the start of the calibration period and the forecasting period (although the author does not discuss this source of uncertainty). To be complete, we should also note that there will be a "structural error", or conceptual error, in the model. The manner in which the model's state variables interact with each other and the forms of the expressions used to describe these interactions will diverge from the (unknowable) "truth". There is currently no adequate method of accounting for errors of this form. This is hardly surprising: the problem is more philosophical than technical.

Response: The above comments provide an excellent description of the sources of uncertainty in the analysis and have been incorporated into the report.

Given the decisions to account for uncertainty in this problem and to account for it using the Kalman filter, lumping the uncertainty in this manner under the single quantity (Σ_Q) is a pragmatic restriction, consistent with benefitting from the relative computational economy of the linear Kalman filter, when set against the alternative of Monte Carlo simulation, say. It also avoids having to use an extended Kalman filter, which would be necessary if one were to separate out (from Σ_Q) the parameteric uncertainty of the model, i.e., item (i) of the sources of uncertainty listed above (see also Beck and Halfon, 1991). The disadvantage of working with such an aggregate measure of uncertainty is that it will foreclose on any analysis of ranking the relative importance of the various sources of uncertainty, in terms of their contributions to the uncertainty of the predictions. Knowledge of this latter would be important in subsequently setting priorities for work that would be needed in order to reduce prediction uncertainty to some acceptable level (if it were thought to be unacceptably high for the purposes of making decisions).

PROPER USE OF THE KALMAN FILTER

As far as can be determined, Yearsley has used the Kalman filter -- for the purpose of calibration -- strictly in accordance with the procedures set out in the paper by van Geer *et al.* (1991). These authors, in their turn, make reference to the covariance-matching procedure of Mehra (1972), which, in spite of its vintage, remains the most common method for calibrating the variance-covariance properties of the system noise, i.e., for assigning values to the elements of the matrix Σ_Q. For the purpose of predicting the consequences of the policy options, again the filter has been applied in a manner consistent with normal practice (Beck and Halfon, 1991). To this extent, no fault can be found with the filter's application here; technically, the analysis appears to be sound.

There are, however, a number of places where caution should be exercised in interpreting the results of the Report. These are as follows.

(i) Figures 6 through 13 show comparisons of the simulated and observed water temperatures. Although we cannot be certain, it is quite probable that the corrected, or updated, estimate of the water temperature, i.e., $\underline{T}_k(+)$ from equation (9), has been used as the "simulated" value. If this is so, it is important to bear in mind that the results of these Figures may suggest a performance of the model better than what would have been achieved in the more familiar, purely deterministic setting, wherein the model is not embedded within a Kalman filter. Close inspection of equation (9) reveals the presence of the current observation of temperature z_k. The effect of updating the one-step-ahead prediction, $\underline{T}_k(-)$, is thus always to draw any erroneous such prediction back towards the observation. The updated estimates $\underline{T}_k(+)$ reflect the benefit of this correction. To the eye trained on assessing a model's performance in the deterministic setting, without this "tracking" feature, a comparison of $\underline{T}_k(+)$ with the observation z_k can be deceptive. It might therefore be desirable to ask for clarification of whether the "simulated" values of Figures 6-13 represent $\underline{T}_k(+)$ or $\underline{T}_k(-)$.

Response: The report has been changed to clarify the way in which simulated results are compared to observations.

(ii) Of the three policy options assessed (business-as-usual, removal of dams, control of tributary temperatures), the removal of the dams will clearly lead to a hydraulic regime unlike that of the (post-dam) observed record. The most obvious expectation of the consequences of this is that the uncertainty attaching to the hydraulic parameters estimated through the approximations of equations (13) through (15), if not any of the other (deterministic) model parameters, will be greater for this regime than for the presently observed conditions (with the dams in place). As far as can be established, no account is taken of this greater uncertainty; the same values of Σ_Q are used in generating the confidence bounds around all three sets of predictions. Since removal of the dams -- on the basis of the current analysis -- is predicted to have a significantly beneficial impact on lowering the number and magnitude of violations of the maximum temperature constraint, more detailed consideration of this point may well be warranted. Furthermore, the potential significance of this particular source of uncertainty may make it appropriate for future analyses to be based on explicit representation of the constituent sources of uncertainty, as opposed to their being lumped under Σ_Q.

Response: Although the variance of the systems model error, Σ_Q, is the same for all scenarios, the variance of the state estimates for the prediction mode of the filter are not necessarily the same. The systems dynamics play an important role in estimating the variance and, in fact, the estimated variance for the scenario in which dams are removed are generally larger than those for the scenarios with the dams in place. It is also important to point out that policy-making is not an objective of this report. The three scenarios were designed to assess the relative importance of dams and tributaries in the thermal regime of the Columbia and Snake rivers.

(iii) It appears that the variance-covariance matrix of the system noise (Σ_Q) has non-zero elements on its leading diagonal alone, i.e., the assumption has been made that disturbances of the stream temperature dynamics are uncorrelated (primarily in space, it would appear). The Report is largely silent on the making of this assumption, although it is a common and not unreasonable one. Nevertheless, there is no discussion of its possible consequences, which is unfortunate since these may be material to the analysis. It is fairly widely appreciated that covariance among the elementary sources of uncertainty can have a significant effect on the propagation of uncertainty. In fact, it has generally been thought that it has the effect of reducing the degree of uncertainty attaching to the forecasts (this is not always the case, however; Beck and Halfon, 1991). We may note that van Geer *et al.* (1991) provide a means of assigning values to these off-diagonal elements of Σ_Q; it does not appear to have been used in the present analysis.

Response: As described in the report, implementation of the mixed Lagrangian-Eulerian approach simplifies the filtering problem since the variance-covariance matrix, $P_k(-)$, has only one element. This eliminates the need to make assumptions regarding the off-diagonal terms. However, it does imply there may be numerical dispersion of the estimated variance as a result of interpolation at the segment boundaries.

(iv) Comparing Figures 14 through 21 with respectively Figures 6 through 13 of the Report, is a surprisingly confusing task. If the principal issue at stake in this study is under-prediction of the maximum (summer) temperatures, it is especially important to be comfortable with the fact that the innovations ($\underline{\nu}_k$) are consistent with the relative positions of the quantities, $\underline{T}_k(+)$ (assumed) and $\underline{z}_k$, plotted in their respective Figures. Even after considerable reflection, I have failed to reconcile -- to my satisfaction -- the two sets of Figures.

Response: The report has been changed in an effort to clarify issues related to this comment.

To summarize, the subject of this review is a Report on a screening analysis designed to identify issues for further study. In general, the Kalman filter has been properly used for this purpose. However, the author of the Report has not identified all of the issues worthy of more detailed scrutiny.

APPROPRIATENESS OF FILTERING FRAMEWORK

In strategic terms, as already stated, it seems appropriate for uncertainty and risk to be parts of this assessment. In tactical terms, the Kalman filter provides (literally) a first-order approximation of error propagation. On balance this would appear commensurate with a preliminary screening analysis, although it is not common to find the Kalman filter employed in a study of this kind. In general, one could say the filter is often a good technique for problem discovery and definition, but one might subsequently want to apply some other form of analysis of the so defined subsequent problems.

Technically, if further use is to be made of the Kalman filter in assessing the Columbia river problem, it would be desirable to investigate the validity of assuming Gaussian distributions for the measurement errors and other sources of uncertainty. Significantly skewed distributions could compromise interpretation of the robustness of the predicted policy outcomes. Likewise, if (deterministic) parametric uncertainties are to be "unpacked" from the single aggregate (of the matrix Σ_Q), and a filtering framework remains the preferred computational setting, this could be achieved through the relatively minor extension of the extended Kalman filter (as in Beck and Halfon (1991)).

ALTERNATIVES

The obvious alternative to using the Kalman filter on a problem of this nature is Monte Carlo simulation, or some variation on that theme. Had this alternative been adopted, uncertainty would almost certainly have been accounted for in a different manner. In particular, as with virtually all Monte Carlo studies, the uncertainty attaching to the (deterministic) parameters of the model would have been the sole source of uncertainty accounted for. The question for calibration would then have been that of using the past observed temperatures in order to constrain, in some way, the choice of candidate parameterisations to be used for predicting the outcomes of the policy alternatives. Normally, one encounters Monte Carlo simulation in the context of forecasting (not model calibration). This requires specification of the statistical distributions to be used for the model's parameters, treated as random variables. In the absence of past observations, ranges of parameter values drawn from the literature are used to define these distributions. It is unusual to find studies using the set of past observations to generate "posterior" distributions of the parameters, for the purpose of forecasting, with the calibration process started with the "prior", literature-derived distributions.

Response: As stated previously, forecasting water temperatures in the Columbia and Snake rivers was not the objective of this Report. Rather it was to compare the thermal regimes of three different scenarios assuming the 21-year record of hydrology and meteorology for the period 1975-1995 was representative of the external driving forces and that the basic management strategies had not changed during this period.

In short, we derive models from uncertain theories reconciled with uncertain observations; we make predictions that are uncertain using models whose uncertainties will reflect all the successes and failures of calibration; and we must make decisions that are robust in the face of the resulting uncertain predictions, i.e., we must determine whether we would opt for the same course of action, all the uncertainties notwithstanding. Conceptually, the Kalman filter fits well with this view. If the alternative of Monte Carlo simulation were to be considered, it would probably find appropriate implementation through the procedure of Generalised Likelihood Uncertainty Estimation (GLUE) of Beven and Binley (1992).

POSSIBLE ISSUES FOR FURTHER STUDY

In the light of what has just been stated, regarding the account taken of uncertainty, from model development, through calibration and forecasting, into decision-making, the following could be of some significance. If one accepts the suggestion that the critical decision will turn on the reliability of the forecasts of maximum temperatures, then the manner in which the model is calibrated -- as the instrument of making this particular prediction -- should be geared to this goal. In practical terms, this implies that the covariance-matching technique employed for choosing Σ_Q should seek the best possible match over the periods of the summer maxima (as opposed to other seasons of the year, or over the year in some average manner). Figures 22 through 29 of the Report do not fully illuminate whether such a strategy has been pursued. We may probably conclude it has not.

Response: The parameter estimation process was performed without including a seasonal bias.

Two criteria are used separately to rank the three policy alternatives, the number of days during the year when the temperature standard is exceeded and the magnitude of the excess temperature. It may be more meaningful to discriminate on the basis of a composite criterion, designed to capture the sense that the joint action of duration and magnitude of the excess is vital for the well-being of the endangered fish.

The option of removing the dams, in spite of the express consideration of uncertainty, still promises to bring about a significant change in the status quo. This is apparent from Figure 34 (when compared Figure 33) and, marginally more so, from the comparison of Figures 39 and 40. Making decisions under uncertainty -- as opposed to the determinism prevailing in its absence -- introduces greater subtlety (and complexity) into the debate. For example, in another context (Klepper *et al.*, 1991) the consequence of an action was forecast to have the effect of increasing the mean value of a commercial mussel culture, but also of introducing (relative to the status quo) a non-negligible risk of population collapse. While it is apparent that the present Report could have sustained such a more elaborate discussion, none is provided.

Response: Regarding the first sentence of the above comment, it is important to note that the Report is not intended to provide support for making policy regarding the removal of dams from the Columbia and Snake rivers. With respect to the rest of the paragraph, the implied reference to Type II error is noted. However, one can only hope the reference to mussel culture is not meant to lead us to conclude that return to the natural or normative river will increase the risk of population collapse for salmon and steelhead in the Columbia and Snake rivers.

CONCLUSIONS AND RECOMMENDATIONS

This EPA Report, in my opinion, should contribute beneficially to the debate surrounding the survival of endangered species of fish in the Columbia River, precisely because of the way in which it casts its analysis in the setting of uncertainty and risk.

Although an unusual method to use, the Kalman filter has been implemented in a technically sound manner. On the whole the approximations and assumptions made in this implementation are consistent with the style of the investigation, this being that of a screening analysis. By implication, therefore, further study is likely to be needed before decisions on managing the thermal regime of the Columbia and Snake Rivers can be made.

Clarification should be sought on the following points: (i) the precise nature of the "simulated" values plotted in Figures 6 through 13; (ii) the possible impact on the predicted results of the policy alternatives of the likely higher uncertainties attaching to the model's hydraulic parameters in the event of removing the dams; (iii) the possible significance of covariance (as opposed to variance) among the sources of uncertainty accounted for in Σ_Q; and (iv) the consistency of interpretation of the results shown in Figures 14 through 21 relative to Figures 6 through 13.

Response: The Report has been changed in an effort to clarify these issues.

If further study is to be undertaken by the EPA, one should seek to have the following issues addressed (among others raised in this review):

(i) a sensitivity analysis of the influence on prediction uncertainty of (a) an enlarged system noise variance-covariance matrix (Σ_Q), as a consequence of removing the dams, and (b) an altered set of values for the elements of this matrix as a result of gearing its calibration to the goal of matching covariances for the summer temperature maxima;

(ii) an assessment of prediction uncertainty when the specific sources of uncertainty are separated out from the aggregated form of Σ_Q, with a view to ranking the relative importance of these different sources;

(iii) an assessment of the normality of the distributions of various quantities manipulated through the filtering algorithms;

(iv) a more elaborate treatment of the implications of these, and any similar, subsequent, results for the debate surrounding decision-making under uncertainty.

REFERENCES

Beck, M B, and Halfon, E (1991), "Uncertainty, Identifiability and the Propagation of Prediction Errors: A Case Study of Lake Ontario", *J Forecasting*, **10**(1&2), pp 135-161.

Beven, K J, and Binley, A M (1992), "The Future of Distributed Models: Model Calibration and Predictive Uncertainty", *Hydrological Processes*, **6**, pp 279-298.

Klepper, O, Scholten, H, and van der Kamer, J P G (1991), "Prediction Uncertainty in an Ecological Model of the Oosterschelde Estuary", *J Forecasting*, **10**(1&2), pp 191-209.

Mehra, R K (1972), "Approaches to Adaptive Filtering", *IEEE Transactions on Automatic Control*, **10**, pp 693-698.

Van Geer, F C, Te Stroet, C B M, and Zhou, Y (1991), "Using Kalman Filtering to Improve and Quantify the Uncertainty of Numerical Groundwater Simulations. 1. The Role of System Noise and its Calibration", *Water Resources Research*, **27**(8), pp 1987-1994.

M B Beck

1091 Two Oaks Drive
Athens
Georgia 30606-5734